EARLY DETECTION, PREVENTION AND INTERVENTION IN ALLIED HEALTH



International Journal of Allied Health Sciences Vol. 9 SUPP3 (2025)

International Journal of Allied Health Sciences

Vol. 9 SUPP3 (2025)

Copyright Notice

Consent to publish: The Author(s) agree to publish their articles with IIUM Press. Declaration: The Author(s) declare that the article has not been published before in any form that it is not concurrently submitted to another journal for publication, and that it does not infringe on anyone's copyright. The Author(s) holds the IIUM Press and Editors of the journal harmless against all copyright claims.

Transfer of copyright: The Author(s) hereby agree to transfer the copyright of the article to IIUM Press, which shall have the non-exclusive and unlimited right to publish the article in any form, including in electronic media. For the article with more than one author, the corresponding author confirms that he/she is authorized by his/her co-author(s) to grant this transfer of copyright.

The IIUM International Journal of Allied Health Sciences (IJAHS) follows the open access policy.

All articles published open access will be immediately and permanently free for everyone to read, download, copy, and distribute for non-commercial purposes.

IIUM International Journal of Allied Health Sciences at https://journals.iium.edu.my/ijahs/index.php/IJAHS is licensed under a Creative Commons Attribution-Non-commercial 4.0 International License.

EDITORIAL TEAM

PATRON IN CHIEF

Professor Dr. Ahmad Aidil Arafat Bin Dzulkarnain

CHIEF EDITOR

Professor Dr. Mohd. Zulfaezal Bin Che Azemin

EDITOR

Assoc. Prof. Dr. Wan Azdie Mohd Abu Bakar

SECTIONS EDITORS

Professor Dr. Muhammad Muzaffar Ali Khan Khattak
Professor Dr. Suzanah Binti Abdul Rahman

ASSOCIATE EDITORS (TECHNICAL)

Asst. Prof. Dr. Mohamed Arshad Mohamed Sideek
Asst. Prof. Dr. Mohamad Ashraf Rostam
Asst. Prof. Dr. Ummi Farhana Hashim

LANGUAGE EDITOR

Assoc. Prof. Dr. Wan Aslynn Salwani Binti Wan Ahmad
Asst. Prof. Dr. Rozlin Binti Abdul Rahman
Asst. Prof. Dr. Nurulwahida Saad

COPY EDITOR

Asst. Prof. Dr. Norsyuhada Binti Alias

Asst. Prof. Dr. Nuraniza Binti Azahari

SECRETARY

Asst. Prof. Dr. Nor Azwani Binti Mohd Shukri

INTERNATIONAL ADVISORY TEAM MEMBERS

Assoc. Professor Dr. Riyaz M. Basha

University of North Texas Health Science Center 3500 Camp Bowie Blvd, Fort Worth, TX 76107, USA.

Professor Dr. Gilson Khang
South Korea, Biomaterials Lab at Korea Research Institute of Chemical Technology (KRICT, Deajeon Korea).

Professor Dr. Gottipolu Rajarami Reddy
Vikrama Simhapuri University, Nellore-524003 India

Professor Dr. Matcha Bhaskar
Division of Animal Biotechnology, Dept. of Zoology, S.V. University TIRUPATI – 517502

Professor Dr. Rubina Hakeem
Clinical Nutrition Department, Faculty of Applied Medical Sciences, Taibah University, Al Madinah Al Munawaroh, Saudi Arabia Pox 4583 Post Code: 41412.

Professor Dr. Alam Zeb
Faculty of Nutrition Sciences, Agricultural University Peshawar, 25130, Khyber Pakhtunkhwa, Pakistan

Professor Dr. Muhammad Subhan Qureshi
Faculty of Animal Husbandry and Veterinary Sciences, Department of Livestock Management, The University of Agriculture, Peshawar, 25130, Khyber Pakhtunkhwa, Pakistan.

Professor Dr. Rizal Damnik
Department of Community Nutrition, Faculty of Human Ecology, Bogor Agricultural University, Bogor, Indonesia-16680.

ABOUT THE JOURNAL

International Journal of Allied Health Sciences, is a peer-reviewed, English-language scholarly online journal, published biannually by the Kulliyyah of Allied Health Sciences, International Islamic University Malaysia

OBJECTIVES

The main objectives of this journal are to;

- nurture the Allied Health Sciences Professionals in their research dissemination/article writing.
- provide a chance and to review/share knowledge in the related research and professional interest.
- facilitate academics and researchers to elevate their intellectual level interacting through this journal.

AREA OF COVERAGE

This journal will provide an avenue for academics to enhance their intellectual level by reviewing and submitting research articles. This journal broadly covers disciplines namely Nutrition and dietetics, Medical Imaging, Biomedical Sciences, Physiotherapy, Speech and audiology, and Optometry. Furthermore, it covers the sub-disciplines within Nutrition and dietetics (Anthropometric, Biochemical and clinical Nutritional Status Assessments, etc.), in Medical Imaging (Radiographic Techniques, Body, Breast, Musculoskeletal, Cardiovascular, and Paediatrics Imaging), Biomedical Sciences (Biochemistry, Bioinformatics, Immunology, Biomedical Engineering, Biophysics, Biotechnology, Cell Biology, Embryology, Endocrinology, Genetics, Medicinal/Pharmaceutical Chemistry, Microbiology, Parasitology, Pharmacology, Physiology, and Toxicology, etc) Physiotherapy (Rehabilitation, Physical Therapy, and Physiotherapy, etc), Speech & Audiology(Clinical Audiology, Educational Speech-Language Pathology, and Speech Therapy, etc) and Optometry (Clinical, Industrial Optometry and issues on Optometry, etc).

MODE OF PUBLICATION

International Journal of Allied Health Sciences (IJAHS) is published biannually with special issues depending on conferences etc. It is only available online in PDF format.

MANUSCRIPT SUBMISSION

As part of the submission process, authors are required to check off their submission's compliance with all of the following items, and submissions may be returned to authors who do not adhere to these guidelines.

- The submission has not been previously published, nor is it before another journal for consideration (or an explanation has been provided in Comments to the Editor).
- The submission file is in OpenOffice, Microsoft Word, RTF, or WordPerfect document file format.
- Where available, URLs for the references have been provided.
- The text is single-spaced; uses a 12-point font; employs italics, rather than underlining (except with URL addresses); and all illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end.
- The text adheres to the stylistic and bibliographic requirements outlined in the Author Guidelines, which are found in About the Journal.
- If submitting to a peer-reviewed section of the journal, the instructions in Ensuring a Blind Review have been followed.

PUBLICATION ETHICS

The authors are expected to properly quote and give full credibility to the source. Furthermore, take written permissions where the photograph and figures used might be the source/reason for a violation of copyrights. The prospective authors are advised to visit the blog on ethics of the Committee on Publication Ethics (COPE) (http://www.publicationethics.org/)

PUBLICATION POLICY

The International Journal of Allied Health Sciences (IJAHS) is an open-access online journal based on the OJS system. All content in it is freely available without charge to the user or his/her institution. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles in this Journal without asking prior permission from the publisher or the author. This is by the BOAI definition of open access.

PUBLISHED BY

IIUM Press

IIUM Press, International Islamic University
Malaysia Jalan Gombak, 53100 Kuala
Lumpur, Malaysia Phone (+603) 6421-5014,
Fax: (+603) 6421-6298

Whilst every effort is made by the publisher and editorial board to see that no inaccurate or misleading data, opinion or statement appears in this Journal, they wish to make it clear that the data and opinions appearing in the articles and advertisement herein are the responsibility of the contributor or advertiser concerned. Accordingly, the publisher and the editorial committee accept no liability whatsoever for the consequence of any such inaccurate or misleading data, opinion, or statement.

International Journal of Allied Health Sciences

Volume 9, SUPP3, 2025

TABLE OF CONTENTS

COPYRIGHT.....

EDITORIAL.....

ORIGINAL ARTICLES AND REVIEWS

Evaluation on the Employment Profile and Graduates' Perception on Employability Among Biomedical Science Graduates From 2016-2019 of International Islamic University Malaysia (IIUM)

Muhammad 'Ammar Mat Rohim, Norazsida Ramli, Mardhiah Mohamad 1-6

Genetic Testing in Islamic Perspective: Updates and Challenges – A Narrative Review

Mohd Yusof Mohamad, Ahmad Fahmi Harun @ Ismail, Mohamed Arshad Mohamed Sideek 7-12

The Vicious Cycle of Menopause: A Narrative Review of the Interplay Between Urinary Incontinence, Mood Disturbances, and Sleep

Nur Ain Syuhada Mohd Zaini, Mohd Yusof Mohamad, Mohamed Arshad Mohamed Sideek, Ahmad Fahmi Harun Ismail 13-18

University Students and Excessive Screen Time: Links to Sleep Disturbance and Acute Health Issues

Norhidayah Ahmad, Norfaezah Ahmad, Najiyya Abdullah Kamil 19-28

Intraocular Pressure Changes During Valsalva Manoeuvre in Young and Middle-Aged Adult Populations

Syazwan Hafiz Mohd, Firdaus Yusof 29-35

Trend and Distribution of Refractive Errors Among Children Attending a Private Optometry Practice in Northern Malaysia

Nadiah Rashidi, Muhammad Afzam Shah Abdul Rahim, Firdaus Yusof 36-43

The Effectiveness of Dichoptic Mobile Therapy

Noor Wafirah Shafee, Norsyahirah Mohamad Azmi, Nur Amani Izzati Mat Ghani 44-50

Digital Learning's Toll: A High Prevalence of Computer Vision Syndrome in University Students During Lockdown

Lim Co Shin, Lili Asma Ismail, Shah Farez Othman, Mohd Hafidz Ithnin, Mohamad Hanif Hajar Maidin 51-58

First Words and Expressive Vocabulary Growth in Toddlers: A Cross-Sectional Study

Nur Atiqah Abd Talib, Nur Hanisah Tukiran, W.A Wan Aslynn, Marina L. Alisaputri 59-66

Vegetable Preferences among Preschoolers in Kuantan, Pahang

Nursyafiqah Shaharudin, Pramitha Sari, Nurul Hazirah Jaafar 67-72

Targeting Diabetes Pathways: An in silico Study of Bioactive Compounds from Mitragyna speciosa Leaf Extract

Taslima Begum, Qamar Uddin Ahmed, Mohd Hafiz Arzmi, Syed Adnan Ali Shah, Mohd Salleh Rofiee, Murni Nazira Sarian, ABM Helal Uddin, Muhammad Muzaffar Ali Khan Khattak 73-86

Association Between Field of Study, General Nutrition Knowledge, and Healthy Eating Practice Among IIUM Students

Noraishah Mohamed Nor, Alea Ameera Akmal binti Kama, Muhammad Ashraf Rostam 87-94

Healthy Maternal Nutrition for Child Stunting Prevention and its Promotion Through Animated Video-based Education: A Narrative Review

Amira Farhana Azizan, Muhamad Ashraf Rostam, Mohd Ramadan Ab Hamid, Muhamad Sadry Abu Seman, Rahiwan Nazar Romli, Nor Azwani Mohd Shukri 95-102

Acceptability Of Bread Fortified with Jackfruit (Artocarpus Heterophyllus) Flour: Evaluation of Sensory Qualities

Ibrahim Abu Bakar, Nurulhuda Mahasan 103-111

A Survey on The Acceptance of Non-Nutritive Sweeteners among Hospitalised Diabetic Patients

Nur Wadhihah Abdul Jalil, Roszanadia Rusali, Noor Suzana Osman, Wan Fathin Fariza Wan Mahmood 112-119

Energy and Protein Intake and Its Association with Malnutrition Risk among Community-Living Older Adults in Kuantan, Pahang

Siti Nur Amirah Sheikh Hishamuddin, Nur Izzati Mohammad Rosli, Aliza Haslinda Hamirudin 120-128

Anthropometric Profile, Sports Nutrition Knowledge and Its Association Among Karate Athletes in Kuantan, Pahang

Muhammad Amirul Aiman Muhammad Yuswadi, Norashikin Mustafa 129-139

A Comparison of Knowledge, Attitude and Practices Related to Milk and Dairy Products among Male and Female Orphanage Adolescents in Kuantan, Pahang

Siti Nur Amirah Sheikh Hishamuddin, Nur Syazwani Azmi, Aliza Haslinda Hamirudin 140-148

Assessing Foodservice Quality in University Cafeterias for Health Promotion: An Importance–Performance Analysis Approach

Noor Suzana Osman, Sharifah 'Ainur Mardhiah Syed Ahmad Shuxry, Wan Fathin Fariza Wan Mahmood, Ibrahim Abu Bakar 149-158

Extraction of Local Artocarpus heterophyllus (Jackfruit) Seeds and Its Effects on the Migration of Oral Squamous Cell Carcinoma (OSCC)

Annur Sofea Fitriah Binti Rizbi, Mahfuzah Elmuhammady Binti Abu Bakar, Engku Nasiha Binti Engku Ngah Sayuddin, Muhamad Ashraf Rostam 159-163

The Association Between Refined Carbohydrate Intake, Acne Severity And Quality Of Life Among IIUM Students: A Cross-sectional Study

Nuraniza Binti Azahari, Nurin Najihah Bahrin, Noraishah Mohamed Nor, Aflah Afandi 164-169

Nutritional and Functional Characterization of Malaysian Banana Flours (Pisang Berangan and Pisang Nangka) and Their Application in High-Fiber Biscuit Development

Hanapi Mat Jusoh, Normah Haron, Noraslinda Muhamad Bunnori 170-175

Food Neophobia Among Preschool Children in Kuantan, Pahang

Nurul Hazirah Jaafar, Nurul Syamin Mohamed Kamal, Pramitha Sari 176-183

Evaluating the Association Between Dairy Consumption, Acne Development and Quality of Life Among University Students

Nuraniza Binti Azahari, Adrina Munirah Shariman, Noraishah Mohamed Nor, Aflah Afandi 184-189

Meal Replacement Plans in Managing Adolescent Obesity: A Systematic Literature Review

Noraishah Mohamed Nor, Ainul Mardhiah Irfan, Roszanadia Rusali, Noor Suzana Osman 190-196

Eating Behaviour and Social Media Addiction among Secondary School Students in Kuantan, Pahang

Khalisah Adnan, Roszanadia Rusali, Aliza Haslinda Hamirudin, Wan Fathin Fariza Wan Mahmood 197-203

Knowledge, Attitudes, and Practices of Laboratory Biosafety Among Undergraduate Students in a Malaysian Public University: Lab Biosafety KAP among Malaysia Undergraduates

Syaza Nashitah Shamsol Kamal, Norfazlina Shaharuddin, Hanani Ahmad Yusof 204-214

Maternal Exposure to Polycyclic Aromatic Hydrocarbon (PAH) and Its Effects on Pregnancy Outcomes: A Systematic Review

Norhidayah Binti Ahmad, Nabilah Husna Mohd Nazumi 215-226

Comparison of the Radioactivity Level of Bird's Eye Chilli in Chemical and Natural Fertilizers

Fatihah Syafinaz Kamarul Zaman 227-233

Prevalence and Ergonomic Risk Profiling of Musculoskeletal Pain in Medical Imaging Students: A Basis for Prevention and Intervention.

Liyana, Hanna Yasmin Harun 234-241

Translation and Cultural Adaptation of the Malay Version of the Noise Exposure Questionnaire (M-NEQ)

Noraidah Binti Ismail, Nur Alia Natasha Apandi 242-253

Perceptions of Occupational Noise, Hearing Protection, and Risk Among Airport Workers: A Preliminary Study from Kuantan, Malaysia

Noraidah Ismail, Ili Athirah Roseli, Nurul Syarida Mohd Sakeri 254-261

Validation of Malay Version of the Parents' Feelings and Perceptions Towards Their Child's Speech or Language Disorder (M-PFPSLD) Questionnaire

Masnira Jusoh, Amni Syamimi Mohamad Khairi 262-283

Psychometric Properties of the Malay Version-Revised Five-To-Fifteen (5-15R) Parental Questionnaire (Malay-5-15R)

Masnira Jusoh, Nurul Najwa Md Yusoff 284-305

An Exploratory Corpus Study of Malay Word Structures in School-Aged Children

Fatin Shatirah Fakrul Razi, Saiful Adli Jamaluddin, Wan Aslynn Salwani Wan Ahmad 306-312

Case Reports of Middle Ear Muscular Spasms Presenting as Tinnitus

Nurlin Ali Hanafiah, Saiful Adli Jamaluddin, Nur Airin Zulaikha Zamzuri, Yahia F. Hussien Al-Hadeethi, Wan Najibah Wan Mohamad 313-317

Evaluation of Indoor Bacterial Load and Its Association with Respiratory Symptoms among Children in Kuantan Early Childhood Centres

Noor Fatihah Mohamad Fandi, Nur Aliya Maisarah, Nurul Falah, Norsyuhada Alias, Maryam Zahaba, Nor Ashikin Sopian 318-325

Social Media Addiction and Stress as Predictors of Eating Behaviour Among Undergraduate Students: A Cross-Sectional Study

Muhamad Ariff Ibrahim, Nurulwahida Saad, Mohd Nazir Mohd Nazori, Siti Adibah Waisulqrnai, Hannah Sorfina Azmi 326-333

LeptoSafe: Integrating Digital Health Tools in Early Detection and Prevention of Leptospirosis

Norafiza Zainuddin, Muhammad Amirul Afiq Rosli, Irfan Sahak, Norsaremah Salleh, Samer Zain, Roslaili Khairudin, Amir Faisal Mohd Khairi, Mohd Arifin Kaderi 334-342

Allergy Prevention Mechanisms of Human Milk Proteomes in Infants: A Scoping Review

Radiah Abdul Ghani, Nurul Faqihah Shaharul, Eka Nurhayati, Tengku Norbaya Tengku Azhar 343-349

Antioxidant-Treated Sperm Enhances Fertilization and Early Embryogenesis in Intracytoplasmic Sperm Injection (ICSI) Cycles

Azantee Yazmie Abdul Wahab, Sardiana Sarmidi 350-360

Artificial Intelligence in Oncology for Early Detection and Intervention: Clinical and Operational Insights from Malaysia

Mardhiah Mohammad, Nik Muhammad Alif Mohd Azmi, Nik Muhammad Alif Mohd Azmi 361-373

Development and Validation of a Questionnaire on Knowledge and Attitude Toward Far-Infrared-Emitting Garments, Their Medical Applications and Health Benefits

Nor Saadah Md Azahri, Ibrahim Adham Taib, Haizatul Umiyah 374-387

Traditional Malay Remedies for Women's Health: Transliteration and Scientific Insights from MSS 1796

Mardhiah Mohammad, Ainul Safiyyah Zainur Rijal 388-397

Time-Lapse Imaging (TLI) versus Preimplantation Genetic Testing (PGT): A Scoping Review on Embryo Assessment Strategies

Azantee Yazmie Abdul Wahab, Muhammad Izzat Zhulkefli 398-410

Anticancer Properties of Bee Venom: A Review

Mohammad Syaiful Bahari Abdull Rasad, Siti Nur Farisyah Abdul Rahim 411-420

Examining The Correlation Between Screen Time, Mental Health, and Dry Eye Syndrome Among Undergraduate Students at KAHS

Norzaris Arieana Zaini, Muhammad Afzam Shah Abdul Rahim, Firdaus Yusof, Ilyanoon Zahari, Noor Haziq Saliman 421-427

The Effects of Water Consumption on Biometry in Myopic Young Adults

Amirah Irdina As'ari, Nureen Syaza Che Mohd Zain, Muhammad Afzam Shah Abdul Rahim, Mizhanim Mohamad Shahimin, Sabrina Subri, Firdaus Yusof@Alias 428-434

Comparative Evaluation of Honey, Aloe Vera, and Olive Oil as Alternative Ultrasound Coupling Agents for Abdominal Imaging

Ummi Farhana Hashim 435-441

Study on Mental Health Status, Possible Influence Factors and Self-Coping Approach among International Islamic University Malaysia (IIUM) Medical Imaging Students

Norhanna Sohaimi, Muhammad Putera Hakimi Mohd Ghazali, Fatihah Syafinaz Kamarul Zaman, Umi Nadrah Amran 442-449

Utilising MRI Data and VOI-Based Analysis for Enhanced Epilepsy Prediction: A Translational Approach from Bench to Bedside

Nur Arif Che Mat, Nur Nadiah Mohd Rais, Mohd Zulfaezal Che Azemin, Intan Bazilah Abu Bakar, Iqbal Jamaludin 450-456

Investigating the Association Between Body Mass Index (BMI) and Body Image Perceptions Among High School Students

Nur Fadhlin Najwa Rosman, Roszanadia Rusali, Aliza Haslinda Hamirudin, Ibrahim Abu Bakar 457-464

Understanding Supervisors' Perspective on Clinical Competencies and Challenges in Dysphagia Training for Speech-Language Therapy Students

Natrah Ahmad Nordin, Nor Azrita Mohamed Zain, Ahmad Aidil Arafat Dzulkarnain, Norsila Abdul Wahab 465-477

Cognitive Interference Effects of Colour-Word and Counting Stroop Task on the Auditory Brainstem Response (ABR) in Children and Adults

Nur Hafizah Sulaiman, Muhamad Kamarul Asyraf Bakri, Muhammad Imran Yusri, Rina Kartini Alhawari, Ahmad Aidil Arafat Dzulkarnain 478-487

Virtual Reality-Based Vestibular Rehabilitation Therapy: A Scoping Review

Nurul Syarida Mohd Sakeri, Saiful Adli Jamaluddin, Megat Ikmal Megat Abdul Karim @ Rosmady 488-495

Analysis of Auditory Brainstem Response test time using Narrow Band Level Specific CE-Chirp and Tone Burst in Infants

Ahmad Aidil Bin Dzulkarnain, Amnah Mohamad Noor, Saiful Adli Jamaluddin Jamaluddin, Sarah Rahmat, Siti Rashidah Mohd Zain 496-502

The Impact of Shared Book Reading on the Speech and Language Development of Infants and Toddlers Aged 0-36 MonthsA Scoping Review

Juliana Aminah Marhaban, Aishah Humaira', Mimi Salleh 503-522

Anxiety Level Among Parents of Newly Diagnosed Children with Speech-Language Impairment: A Study Using the Malay Version of DASS

Juliana Aminah Marhaban, Nuratirah Nazihah Ngadimin, Rosnani Sarkarsi 523-529

A Scoping Review of Clinical Practice Guidelines for the Assessment of Sarcopenia

Siti Salwana Kamsan, Aliza Haslinda Hamirudin, Nor Syuhaidah Udin 530-543

The Comparison of Contact Lens Wear Clinical Symptoms Between Different Modality of Lenses

Nur Amani Izzati binti Mat Ghani, Noor Ezailina binti Badarudin 544-550

ORIGINAL ARTICLE

Evaluation on The Employment Profile and Graduates' Perception on Employability Among Biomedical Science Graduates from 2016-2019 of International Islamic University Malaysia (IIUM)

Muhammad 'Ammar Mat Rohim, Norazsida Ramli*, Mardhiah Mohamad

Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

ABSTRACT

Background: The employment rate among fresh graduates in Malaysia, particularly Biomedical Science graduates, has become a major concern. This survey was conducted to assist the Biomedical Science Education Provider (BSEP) at the International Islamic University Malaysia (IIUM) in tracking alumni employment profiles and gathering feedback for improving learning and teaching activities. The insights gained from this survey will enable the Academic Education Provider (AEP) to evaluate educational outcomes, contributing to the continuous enhancement of biomedical science education standards. **Methods**: An online questionnaire was used as the primary tool for data collection. The questionnaire assessed graduates' job profiles, perceptions of how their education and training aligned with their current employment, and included a section for additional comments and suggestions. A total of 92 graduates (38.3% response rate) from a pool of 240 Biomedical Science alumni participated in this study. **Results**: Findings indicated that 69.6% (n=64) of respondents were employed, 10.9% (n=10) were unemployed, and 19.6% (n=18) were outside the labour force, with 17 of these pursuing further studies. A significant association was found between curriculum score and employment status (p=0.029, Phi=0.317). However, no significant associations were observed between academic performance or year of graduation and employment status. **Conclusions**: In conclusion, the majority of IIUM Biomedical Science graduates benefited from the knowledge and training provided during their studies. The comments and suggestions collected through this survey will be used to inform and strengthen the next curriculum review.

Keywords:
employment status; Biomedical Science; employed; unemployed; outside labour force

INTRODUCTION

The combined forces of globalisation and rapid technological advancement have intensified competition in the job market, placing increasing pressure on graduates seeking employment. In Malaysia, this challenge has been compounded by the economic disruptions caused by the COVID-19 pandemic. The Department of Statistics Malaysia (DOSM, 2020) reported that national economic growth for the first quarter of 2020 fell to 0.7%, a significant decline from 3.6% in the previous quarter and the lowest recorded growth since the third quarter of 2009. Based on analyses aligned with international standards developed by the United Nations, World Bank and International Monetary Fund, DOSM estimated that the economy would have expanded by 3.9–4.2% had it not been impacted by the pandemic. Instead, Malaysia suffered an economic output loss of approximately RM22.8 billion in the first quarter alone (DOSM, 2020). This downturn has intensified job competition, particularly among recent graduates and individuals displaced from the workforce during the pandemic. In the health and welfare disciplines, including biomedical science, 6,742 students graduated from Malaysian public universities in 2019; however, the domestic job market lacks the capacity to absorb this annual influx of graduates (MOHE, 2020).

Within this broader economic context, the employability of biomedical science graduates faces additional discipline-specific challenges. The field is undergoing rapid transformation, driven by the rise of large-scale, data-intensive and interdisciplinary research that increasingly demands proficiency in bioinformatics, functional genomics, computational analysis, and collaborative scientific practice (National Research Council, 2011;

* *Corresponding author.*
E-mail address: norazsida@iium.edu.my

Lenfant, 2000). Yet many undergraduate programmes remain rooted in traditional laboratory training, resulting in a skills mismatch that undermines graduates' readiness for modern biomedical and diagnostic environments. Employers continue to highlight gaps in practical competencies, digital literacy and workplace preparedness, reflecting a persistent disconnect between academic instruction and real-world expectations (Scott et al., 2024). These challenges are further compounded by the expanding complexity of biomedical science, which often necessitates extended or specialised training in areas such as systems biology, translational research and point-of-care diagnostics (Lenfant, 2000). Additionally, inconsistent incorporation of employability focused components such as simulation-based learning, interprofessional training and competency-based assessments across institutions contributes to variability in graduate readiness and competitiveness (Scott et al., 2024). Addressing these issues requires a more systematic integration of interdisciplinary, data-driven and professional skills within biomedical science curricula to ensure graduates are adequately equipped for an evolving and demanding workforce landscape.

MATERIALS AND METHODS

Study Design and Subjects

This study employed a cross-sectional design, as data were collected at a single point in time through an online questionnaire. This design allowed for the simultaneous measurement of exposures (academic performance, curriculum score, and year of graduation) and outcomes (employment status prevalence). Purposive sampling was used to focus specifically on Biomedical Science graduates from the International Islamic University Malaysia (IIUM) between 2016 and 2019, prior to the Covid-19 pandemic that significantly disrupted the global economy and resulted in approximately 13 million people becoming unemployed worldwide (Verma & Prakash, 2020). Inclusion criteria required participants to be former Biomedical Science students who graduated from IIUM between these years, while exclusion criteria included students who did not complete their studies or were terminated. The total target population comprised 240 graduates from 2016 to 2019. Out of these, 92 graduates responded, resulting in a response rate of 38.2%.

Questionnaire

The questionnaire was adapted from the tracer study surveys conducted by the Ministry of Higher Education (MOHE). The questionnaire consisted of three sections: Part A gathered demographic information; Part B recorded current job status (employed, unemployed, outside the labor force); and Part C assessed graduates' perceptions on employability, specifically focusing on suggestions for enhancing the marketability of Biomedical Science graduates and identifying preferred training programs.

Statistical Analysis

Data were analyzed using SPSS software version 25.0. Descriptive statistics were used to determine the prevalence of employment profiles among IIUM Biomedical Science graduates. Inferential statistics, including the Chi-square test, were applied to examine associations between categorical variables within the same population.

RESULTS

Demographic Profile of Graduates

Out of the 240 graduates from 2016 to 2019, the study received 92 responses, resulting in a response rate of 38.2%. A post hoc power analysis for the chi-square test of independence (df = 1, $\alpha = .05$, Cohen's $w = 0.30$) with the obtained sample size of N = 92 indicated a statistical power of approximately 0.82, which is generally considered adequate. Female graduates outnumbered male graduates, comprising 71.7% (n=66) female and 28.3% (n=26) male respondents. Participants' ages ranged from 25 to 30 years. Table 1 presents the response rate for each graduation year, with the highest response from the 2019 cohort (49.1%, n=28), followed by the 2018, 2017, and 2016 cohorts. All participants were Malaysian nationals, with most residing in various states across Malaysia, except for one graduate currently residing in the United Kingdom for further studies.

Table 1: Total Responses of graduates based on each year of graduation

Year of Graduation	Total graduates	Total graduates participated in the study	Percentage (%)
2016	53	13	24.5
2017	70	24	34.3
2018	60	27	45
2019	57	28	49.1

Academic Performance and Curriculum Score of Graduates

Academic performance was categorized by degree class. Table 2 shows that 15 graduates achieved a first-class degree, 75 graduated with an upper-second class degree, and two with a lower-second class degree. A chi-square analysis indicated a significant association between gender and academic performance, χ^2 (2, n=92) = 6.768, p=0.034, Cramer's V = 0.271, showing a small association, with female graduates tending to perform better academically than their male counterparts. Graduates also self-assessed

their curriculum involvement, scoring from one (not active at all) to five (very active). Seven graduates rated below average, while 85 scored above three. The median curriculum score of four indicated that the majority of graduates were actively engaged in extracurricular activities.

Table 2: Academic performance of graduates based on class of degree

Class of Degree	Frequency	Percentage (%)
First class (3.50 and above)	15	16.3
Upper-Second class (2.80-3.49)	75	81.5
Lower-Second class (2.00-2.79)	2	2.2

Employment Status of Biomedical Science Graduates

Employment status was classified into three categories: employed, unemployed, and outside the labour force. Employed refers to individuals who are currently engaged in any form of paid job, whether full time, part time, contract-based, or self-employed. Unemployed refers to individuals who are not working but are actively seeking and available for employment. Outside the labour force refers to individuals who are neither working nor actively seeking work, such as those pursuing further studies or those not currently available for employment. Of the graduates, 64 were employed, 10 unemployed, and 18 outside the labour force, with 17 of these pursuing further studies. The chi-square test results showed no significant association between the year of graduation and employment status, $\chi^2(3, n=74) = 3.816$, $p=0.282$, indicating that graduation year did not influence employment outcomes. Likewise, no significant association was found between academic performance and employment status, $\chi^2(2, n=74) = 0.526$, $p=0.769$, suggesting that academic performance alone did not impact employment status. However, a significant association was found between curriculum activity scores and employment status, $p=0.029$, Phi=0.317, highlighting that graduates actively engaged in extracurricular activities had higher employability.

Graduates' Perceptions on Employability

In the employability perception section, graduates were presented with statements on factors affecting employment, including academic performance, skills, and education level, and asked to agree or disagree. Most graduates (95.7%, n=88) agreed that "both academic performance and skills are important for securing a job." Additionally, 82 graduates believed that "academic performance, skills, and a higher education level are critical for employment."

An open-ended question invited graduates to suggest ways to enhance the employability of Biomedical Science graduates. Graduates also indicated the types of training they would like to pursue to improve their job prospects. A total of 33 graduates (35.9%) recommended increasing practical sessions over theoretical classes. Fourteen graduates suggested that students should focus on developing soft skills such as communication, interpersonal abilities, and public speaking through extracurricular activities. Twelve graduates recommended that students gain job exposure in the biomedical science sector to familiarize themselves with potential career paths. Among the types of training preferred, 35.9% (n=33) indicated interest in laboratory certifications, such as polymerase chain reaction (PCR) training, quality laboratory management, and phlebotomy. Additionally, 31 graduates expressed a desire to attend soft skills training, including leadership and communication, to enhance employability.

DISCUSSION

Demographic Profile of Graduates

Among the 92 responses, female graduates exceeded male graduates. The DOSM reported that female enrolment at the tertiary level surpassed that of males, reaching 50.3% (DOSM, 2019), resulting in more female graduates annually. However, this distribution varies by field; data from the MOE indicate that, in 2018 and 2019, female graduates outnumbered male graduates across all fields except for "Engineering, Manufacturing, and Construction," where male graduates were predominant (MOHE, 2019).

According to Hdii & Fagroud (2018), female students often outperform males academically, which may explain the observed association between gender and academic performance in this study, with most female graduates achieving higher academic results than males. The median curriculum activity score of four (active) aligns with the 2019 IIUM tracer study report, which found that most IIUM students were highly involved in co-curricular activities (Selamah et al., 2018). This result affirms that IIUM successfully cultivates students who perform well academically and actively participate in extracurricular activities.

Employment Status of Biomedical Science Graduates

The MOHE tracer study data indicated that the employment rates of IIUM Biomedical Science graduates from 2016 to 2019 were 69.8%, 55.7%, 66.7%, and 77.4%, respectively, with 2019 recording the highest pre-convocation employment rate (Ministry of Higher Education, 2019). This translates to an average

employment rate of 67.4% for this period. Current study data show that the overall employment rate for graduates from 2016 to 2019 has risen to 69.6% (n=64). According to DOSM's 2019 employment statistics, the trend observed in this study where employed graduates having the highest percentage, followed by those outside the labour force and then unemployed graduates, mirrors the national trend (DOSM, 2019).

Although no significant association was found between year of graduation and employment status, an imbalance in response rates across graduation years could have influenced this analysis. It was expected that the 2016 cohort, with over three years since graduation, would have the highest employment rate due to longer time and greater experience in the job market. This study also found that academic performance alone did not significantly impact employment status. As noted by Tentama and Abdillah (2019), while higher academic achievement may enhance employability, skills are equally important alongside strong academic performance. The significant association observed between curriculum activity scores and employment status suggests that skills, particularly those developed through extracurricular involvement, play a critical role in employment outcomes. This finding aligns with studies indicating that active involvement in co-curricular activities positively influences graduate employability (Akinrinmade & Ayeni, 2017).

The unemployment rate of 10.9% among IIUM Biomedical Science graduates prior to the Covid-19 pandemic can be explained by several structural and graduate-related factors. Employment opportunities in core biomedical fields are relatively limited and competitive, often favouring candidates with specialised laboratory experience or prior exposure to regulated environments, a pattern also noted in STEM-related labour-market analyses (Rahmah, Ishak, & Lai, 2011). Graduates additionally compete with applicants from closely related science disciplines, which further intensifies competition for entry-level positions. Moreover, some graduates delay entering the workforce due to personal commitments or preparation for postgraduate studies, placing them temporarily in the unemployed category. Location constraints, strong preference for public sector positions, and the generally longer transition period observed among science graduates in Malaysia (Ministry of Higher Education, 2020) also contribute to the overall unemployment rate observed in this cohort.

On the other hand, the substantial number of IIUM Biomedical Science graduates who pursued further studies reflects the academic and professional nature of the field, which often requires advanced qualifications for meaningful career progression. Biomedical Science is heavily research-oriented, and many roles within medical research, diagnostics, and laboratory-based sciences prefer candidates with postgraduate training due to the need for specialised methodological skills and research competency. These graduates may also view postgraduate education as a strategic pathway to enhance employability and long-term career prospects, a trend that aligns with national patterns in STEM disciplines where graduates frequently extend their studies to strengthen their professional competitiveness (Rahmah et al., 2011; Ministry of Higher Education, 2020). Furthermore, strong interest in scientific discovery and encouragement from academic mentors likely contributed to their decision, making postgraduate studies a natural continuation of their academic trajectory.

Graduates' Perception on Employability

Understanding graduates' perceptions of employability is essential to bridge the gap between employer and graduate expectations. Discrepancies between these perspectives can contribute to the rising number of unemployed graduates (Hwang, 2017). Most graduates agreed that both academic performance and skills are crucial for securing employment. While academic performance is essential, technical and employability skills have become increasingly important in today's competitive job market (Mwita, 2014). Many graduates recommended increasing practical sessions over theoretical classes. Although IIUM provides sufficient lectures and practical sessions, graduates believe additional hands-on experience would enhance technical skills, thereby improving employment prospects (Mwita, 2024).

Graduates also emphasized the importance of soft skill development, noting that skills such as communication, personal qualities, teamwork, problem-solving, organizational skills, technological proficiency, and continuous learning all influence employability (Abdullah et al., 2019). They further suggested that students should gain job exposure, recommending that the department invite alumni or industry representatives to share job-related insights and experiences relevant to Biomedical Science graduates.

Graduates identified specific training they believe would be beneficial, including laboratory certifications (e.g., PCR training, quality laboratory management, and phlebotomy), which could enhance their qualifications for roles in the healthcare sector. Given the rapid development of new technology, adequate training on advanced tools and techniques is essential for job performance (Roco et al., 2002). Many graduates also

chose soft skills training, as employers increasingly value interpersonal skills and professional expertise as essential for job security and advancement (Sato et al., 2019). Overall, graduates believe that these training programs could aid in securing employment, and they suggest that the university consider implementing these initiatives to better prepare students for post-graduation success.

CONCLUSIONS

In conclusion, the employment outcomes of IIUM Biomedical Science graduates from 2016 to 2019 indicate that 69.6% (n = 64) were employed, 10.9% (n = 10) were unemployed, and 19.6% (n = 18) were outside the labour force, with the majority of the latter group pursuing further studies. The analysis identified curriculum score as a significant factor influencing graduates' employment status, whereas academic performance and year of graduation showed no meaningful association. These findings suggest that the Biomedical Science programme has largely succeeded in producing graduates who are competitive and equipped with relevant knowledge and skills for the job market.

However, the study is not without limitations. The total number of respondents was relatively low, which may restrict the generalisability of the findings to the broader population of IIUM Biomedical Science graduates. Time constraints also limited the opportunity to extend data collection and increase the sample size. Additionally, the study did not conduct separate reliability and validity assessments for the questionnaire, as it utilised an instrument developed by the Ministry of Higher Education for Malaysian graduates across institutions. While the MOHE questionnaire is widely adopted, the absence of instrument validation within the specific context of IIUM Biomedical Science graduates may introduce potential measurement limitations. Despite these constraints, the study provides valuable insight into graduate employability trends and highlights areas for further investigation.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Abdullah, A. R., Muhammad, M. Z., & Nasir, N. A. M. (2019). The role of soft skills on business graduates employability. *Journal of Entrepreneurship & Business*, *7*(2). https://dx.doi.org/10.17687/JEB.0702.07

Akinrinmade, B. I., & Ayeni, A. O. (2017). Influence of extracurricular involvement on graduate employability. *MOJEM: Malaysian Online Journal of Educational Management*, *5*(3), 19-31. https://doi.org/10.22452/mojem.vol5no3.2

Department of Statistics Malaysia. (2019). *Women empowerment in Malaysia 2019.* Department of Statistics Malaysia. https://www.dosm.gov.my

Department of Statistics Malaysia. (2020). *Malaysia economic performance first quarter 2020.* Department of Statistics Malaysia. https://www.dosm.gov.my

Hdii, S., & Fagroud, M. (2018). The effect of gender on university students' school performance: the case of the National School of Agriculture in Meknes, Morocco. *Culture & Society*, *9*(1), 67-78. https://dx.doi.org/10.7220/2335-8777.9.1.4

Hwang, J. (2017). *Understanding the employment gap: A study of employer and graduate perspectives.* Journal of Career Development, 44(3), 1-12.

Lenfant, C. (2000). Training the next generation of biomedical researchers: challenges an oopportunities. *Circulation*, *102*(4), 368-370.

Ministry of Higher Education (2019). Statistik_Pendidikan_Tinggi_2018_-_Bab_7. *Tracer Study, MOE*, 161–171.

Ministry of Higher Education (2020). Statistik Pendidikan Tinggi 2019 : Kementerian Pengajian Tinggi | 47. *Kementerian Pengajian Tinggi Malaysia*, 47–59

Mwita, K. (2024). The role of soft skills, technical skills and academic performance on graduate employability. *Technical Skills and Academic Performance on Graduate Employability (January 30, 2024)*. https://dx.doi.org/10.2139/ssrn.4931628

National Research Council. (2003). New opportunities, new challenges: The changing nature of biomedical science. In *Enhancing the vitality of the national institutes of health: organizational change to meet new challenges*. National Academies Press (US).

Rahmah, I., Ishak, Y., & Lai, W. S. (2011). *Employability of Malaysian graduates: Graduates' attributes, employers' expectations, and the role of higher learning institutions.* Universiti Kebangsaan Malaysia.

Roco, M., Bainbridge, W. Converging Technologies for Improving Human Performance: Integrating From the Nanoscale. *Journal of Nanoparticle Research* 4, 281–295 (2002). https://doi.org/10.1023/A:1021152023349

Sato, K., Nakamuro, M., & Ōwan, H. (2019). *The effect of interpersonal skills on worker performance*. RIETI. https://www.rieti.go.jp/jp/publications/dp/19e045.pdf

Selamah, A. Y., Mohd Nahar, M. A., Izyani, Z., Norazsida, R., & Muhamad Karimi, S. (2018). *Tracer Study Report 2018*. Alumni Realtions Division.

Scott, S., Millar, B. C., & McClean, S. (2024). Education and Training in Biomedical Science. *British Journal of Biomedical Science*, *81*, 13598.

Tentama, F., & Abdillah, M. H. (2019). Student employability examined from academic achievement and self-concept. *International Journal of Evaluation and Research in Education*, *8*(2), 243–248. https://doi.org/10.11591/ijere.v8i2.18128

Verma, A. K., & Prakash, S. (2020). Impact of covid-19 on environment and society. *Journal of Global Biosciences*, *9*(5), 7352-7363.

REVIEW ARTICLE

Genetic Testing in Islamic Perspective: Updates and Challenges – A Narrative Review

Mohd Yusof bin Mohamad[1,*], Ahmad Fahmi bin Harun @ Ismail[1], Mohamed Arshad Mohamed Sideek[1]

[1]Department of Physical Rehabilitation Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Bandar Indera Mahkota 25200, Kuantan, Pahang, Malaysia

Keywords:
genetic testing; Islamic perspective; healthcare.

ABSTRACT

Background: Genetic testing is a rapidly evolving field with significant implications for healthcare, particularly in Muslim-majority societies where religious, ethical, and cultural values deeply influence medical practices. The high prevalence of hereditary disorders underscores the urgency of aligning genetic advancements with Islamic bioethical and legal principles. This review explores genetic testing through an Islamic lens, aiming to present updated insights, examine ethical and social challenges, and identify future directions for policy and research. **Methods:** A narrative review methodology was employed drawing from databases such as Semantic Scholar and PubMed. The search strategy targeted Islamic perspectives on genetic testing, focusing on ethics, legal norms, clinical applications, and policy evolution. From 1,032 initial records, 777 were screened, and 251 met the inclusion criteria, with the 50 most relevant articles included in the final analysis. Only peer-reviewed papers discussing human genetics within an Islamic context were considered, excluding those unrelated to religious or ethical dimensions. **Results:** Islamic bioethics grounded in the Qur'an, Hadith, and jurisprudential principles, supports genetic testing for disease prevention and treatment but raises concerns over non-therapeutic applications. Clinical applications like premarital screening and prenatal diagnosis are expanding in Muslim societies, supported by state policies yet met with varied public acceptance due to differing interpretations of religious permissibility. Emerging technologies such as CRISPR have intensified ethical debates, particularly around germline modification and enhancement. Regulatory frameworks vary widely across Islamic countries, ranging from formal fatwas to national guidelines, often lacking cohesion. Issues of informed consent, confidentiality, and culturally competent genetic counselling remain pressing concerns. Key research gaps include limited empirical studies on Muslim minorities in non-Muslim countries, the psychological impact of genetic testing, and the operationalisation of Islamic ethical frameworks in clinical practice. **Conclusion:** Genetic testing in Islamic contexts presents both opportunities and ethical dilemmas. While disease-related applications gain broad acceptance, newer technologies require cautious deliberation. A harmonised, faith-aligned regulatory approach, bolstered by community engagement and culturally sensitive counselling, is essential for ethically responsible integration of genetics into healthcare in Muslim societies.

INTRODUCTION

Genetic testing and genomic technologies have significantly transformed modern medicine, offering powerful tools for the prevention, diagnosis, and management of hereditary diseases. In Muslim-majority societies, the integration of these innovations into clinical and legal practices must be carefully evaluated through the lens of Islamic bioethics, which is rooted in the Qur'an, the Hadith, and the principles of Islamic jurisprudence (Aqeel, 2007; Shabana, 2022). Applications such as preimplantation genetic testing (PGT), premarital screening, and the use of genetic evidence in legal proceedings raise important questions concerning the sanctity of lineage, the protection of life, and the preservation of individual and familial rights (Arijo, 2020; Shoaib, 2024).

The emergence of advanced techniques including polygenic testing and embryo selection for non-medical traits such as intelligence or physical characteristics has intensified ethical scrutiny. Although some procedures may initially appear to align with Islamic teachings, deeper analysis often reveals concerns related to intention (*niyyah*), certainty (*yaqin*), harm prevention (*darar*), and societal norms (*'urf*) (Chin et al., 2023; Shabana, 2022). These debates are further complicated by the rapid pace of biotechnological development, diversity in scholarly

* *Corresponding author.*
E-mail address: yusofkahs@iium.edu.my

interpretations, and the need for localised ethical guidance that respects both religious doctrine and community well-being (Amalia & Hilmi, 2024; Shanti et al., 2015; Muhsin et al., 2023; Sachedina, 2020). Against this backdrop, clarifying why genetic testing requires careful examination from an Islamic perspective becomes essential. Many Muslim-majority countries such as Saudi Arabia, Malaysia, Iran, Pakistan, and those in the Gulf are expanding national genomic screening initiatives, driven by factors including high rates of consanguinity, public health priorities, and developing legal and regulatory structures (Alkuraya, 2014; Tadmouri et al., 2009). At the same time, Muslim diaspora communities in Europe and North America face their own set of ethical and legal challenges as they engage with biomedical systems shaped by secular norms and differing regulatory expectations (Ghaly, 2012; Padela & Auda, 2020).

As Muslim communities confront the implications of these technologies, a nuanced understanding is required. Especially the one that harmonises scientific progress with religious values. This paper provides an updated examination of genetic testing within Islamic contexts. It explores the application of Islamic ethical principles to emerging technologies, highlights the legal and social challenges associated with their use, and identifies areas where ongoing dialogue, research, and policy development are essential. Through an interdisciplinary approach, the review seeks to contribute to the evolving discourse on how biomedical innovation can be ethically implemented in a manner that remains faithful to Islamic moral and legal traditions.

METHODS

Study Design

This study adopted a narrative review design to examine the ethical, legal, and social implications of genetic testing from an Islamic perspective. The review aimed to consolidate current knowledge, identify thematic patterns, and highlight policy and research gaps relevant to Muslim populations.

Data Sources and Search Strategy

A structured search was conducted using Consensus AI across three primary academic databases: PubMed, Scopus, and Semantic Scholar. Literature published in English between 2000 and 2024 was considered. Search terms included *"genetic testing", "Islamic bioethics", "Shariah", "Muslim ethics", "premarital screening", "genomic policy",* and *"Islamic law and healthcare"*. Additional articles were retrieved through manual searches of bibliographies from key publications and relevant policy reports.

Inclusion and Exclusion Criteria

Studies were eligible for inclusion if they:

- Addressed genetic testing or genomics in relation to Islamic ethics, law, or socio-cultural contexts.
- Discussed clinical, ethical, or regulatory implications relevant to Muslim-majority populations.
- Were peer-reviewed original research, reviews, or recognized policy documents.

Studies were excluded if they:

- Focused solely on laboratory techniques or molecular genetics without reference to Islamic perspectives.
- Did not involve Muslim populations or lacked relevance to Islamic frameworks.
- Were non-English language or lacked full-text access.

Study Selection and Data Extraction

All identified titles and abstracts were screened for relevance as mentioned in Figure 1. A full-text review was performed on potentially eligible articles. A standardised data extraction form was used to collect information on study objectives, methodology, population or context, key findings, and relevance to Islamic bioethical principles.



Figure 1: Flow diagram of literature search using *Consensus AI Tools*

Data Synthesis

A thematic synthesis approach was utilised. Both inductive and deductive coding methods were applied to categorise findings into thematic domains such as:

- Religious and legal frameworks
- Premarital and prenatal testing
- Community perceptions
- Regulatory approaches
- Emerging technologies and ethical challenges

Quality Appraisal

Included studies were appraised for relevance and methodological rigor using adapted qualitative review criteria suitable for ethical and sociocultural literature.

RESULTS AND DISCUSSION

Recent literature highlights both the rapid advancement and the complex challenges of genetic testing within Islamic contexts. The findings in Table 1 synthesises key themes from recent research, focusing on ethical, legal, and social dimensions, as well as community attitudes and policy developments.

Table 1: Key Themes in Genetic Testing from an Islamic Perspective

Theme	Key Findings	Citations
Ethical/Religious Issues	Disease prevention supported; non-medical enhancement controversial	(Chin et al., 2023; Khan & Konje, 2019; Muhsin et al., 2023; Amalia & Hilmi, 2024)
Legal Regulation	Varies by country; trend toward integrating ethics into law	(Novikova, 2019; Al-Aqeel, 2005)
Social Implications	High consanguinity rates; premarital screening growing; concerns about stigma/discrimination	(Shanti et al., 2015; Zhong et al., 2018)
Community Attitudes	Support for health-related testing; hesitancy for controversial uses	(Muhsin et al., 2023; Zhong et al., 2018; Sulaiman, 2018)
Counselling and Education	Need for culturally/religiously sensitive counselling and public education	(Shanti et al., 2015; Zhong et al., 2018)
Emerging Technologies	Therapeutic gene editing cautiously accepted; enhancement and sex selection opposed	(Chin et al., 2023; Muhsin et al., 2023; Amalia & Hilmi, 2024)
Policy/Implementation	Regulatory gaps; need for harmonised guidelines; access issues	(Novikova, 2019; Zhong et al., 2018; Al-Aqeel, 2005)

Ethical, Legal, and Social Issues

Islamic bioethics generally supports genetic testing for disease prevention and treatment, provided it aligns with Shariah principles such as the preservation of life and lineage. However, the use of genetic technologies for non-medical enhancements (e.g., intelligence, physical traits) is widely viewed as inconsistent with Islamic law and raises significant ethical concerns (Chin et al., 2023; Khan & Konje, 2019; Muhsin et al., 2023; Amalia & Hilmi, 2024). A critical ethical concern is informed consent which is foundational to ethical biomedical practice but presents unique challenges in many Muslim-majority contexts. While Islam encourages the pursuit of knowledge and individual accountability, cultural norms such as family-centered decision-making, gender dynamics, and varying levels of health literacy can complicate the process. For example, patients may defer decisions to family elders or male guardians, particularly in more conservative settings, which may limit their autonomy. Furthermore, language barriers, limited access to genetic counselling, and insufficient understanding of complex genomic concepts can hinder truly informed decision-making. Scholars have argued for culturally sensitive consent models that respect

Islamic values while promoting patient autonomy and understanding (Alahmad et al., 2018). In terms of legal regulation, Middle Eastern countries vary in their legal approaches, with some developing national guidelines and others relying on fatwas or local customs. There is a trend toward integrating ethical standards into legal norms, with some countries drafting new laws or updating existing regulations to address genomics and genetic testing (Novikova, 2019; Al-Aqeel, 2005).

As for social Implications, high rates of consanguinity in many Islamic societies increase the prevalence of genetic disorders, making premarital and prenatal genetic screening particularly relevant. For example in Pakistan, β-thalassaemia (β-thal) trait prevalence is estimated at 5–7%, amounting to more than 10 million carriers nationwide. Each year, roughly 5,000 children are newly diagnosed with β-thalassaemia major (β-TM), placing a considerable and continuing burden on affected families and on the national healthcare system (Khaliq, 2022). In the Gulf Cooperation Council (GCC) region, the pattern of haemoglobinopathies is equally concerning: the prevalence of sickle cell disease (SCD) ranges from 0.24% to 5.8%, while the sickle cell trait is reported at rates between 1.02% and 45.8%. Consanguineous marriage remains an important contributing factor, heightening the risk of children being affected by haemoglobinopathies (Abu-Shaheen et al., 2022). Although premarital genetic testing is increasingly accepted and even mandated in some countries, concerns remain about privacy, stigma, and the potential for discrimination (Shanti et al., 2015; Zhong et al., 2018; Sulaiman, 2018).

Community Attitudes and Counselling

Many Muslims express support for genetic testing when it is used for health-related purposes, particularly in the prevention and management of hereditary conditions. However, hesitancy arises when such procedures intersect with religious sensitivities, especially in cases involving pregnancy termination or non-medical sex selection. These attitudes are influenced by a combination of religious beliefs, cultural traditions, and varying levels of genetic literacy across communities (Muhsin et al., 2023; Zhong et al., 2018; Sulaiman, 2018). In this context, genetic counselling plays a vital role in facilitating informed and ethically sound decision-making. Despite its importance, challenges persist in ensuring that counselling practices are culturally appropriate and aligned with Islamic ethical values. Furthermore, there is a pressing need for broader public education to correct misconceptions, enhance awareness, and empower individuals to make decisions that are both medically sound and religiously informed (Shanti et al., 2015; Zhong et al., 2018).

Emerging Technologies and Ongoing Debates

Emerging technologies such as CRISPR gene editing and polygenic embryo selection have sparked significant ethical debate within Islamic discourse. While gene editing is generally considered permissible for therapeutic purposes, Islamic scholars strongly caution against germline modifications and enhancements that may fundamentally alter human nature or contravene the principle of not "changing what Allah has created" (Chin et al., 2023; Amalia & Hilmi, 2024). Similarly, the application of preimplantation genetic testing (PGT) for non-medical sex selection or the pursuit of socially desirable traits, such as intelligence or physical appearance, is widely viewed as problematic. Such practices raise serious concerns regarding gender imbalance, equity, and social justice, and are generally not supported within Islamic bioethical frameworks (Chin et al., 2023; Muhsin et al., 2023).

From the perspective of Islamic bioethics, these discussions rest on core principles such as the preservation of life (*ḥifẓ al-nafs*), the safeguarding of lineage (*ḥifẓ al-nasl*), and the avoidance of harm (*darʾ al-mafāsid*). Within the framework of the *maqāṣid al-sharīʿah*, these principles help guide ethical judgement by emphasising the protection of human dignity, the prevention of harm, and the maintenance of familial and generational integrity. As a result, emerging genetic technologies are considered not only in terms of their potential medical benefits but also through their wider moral, social, and theological consequences.

Policy and Implementation Challenges

The regulatory landscape for genetic testing across Islamic countries is marked by fragmentation, with a lack of unified frameworks leading to inconsistencies in policy development and clinical practice. This regulatory gap has prompted calls for harmonised guidelines that can balance the ethical imperatives of Islamic teachings with the realities of scientific advancement (Novikova, 2019; Al-Aqeel, 2005). In addition to regulatory challenges, issues of access and equity persist, particularly in low- and middle-income Islamic countries where resource limitations, a shortage of trained professionals, and persistent social stigma hinder the widespread availability and acceptance of genetic services (Zhong et al., 2018).

Research Gaps

Despite progress, gaps persist in empirical research on community attitudes, the impact of genetic counselling, and the development of unified regulatory frameworks. There is also limited research on the practical

implementation of Islamic bioethical principles in clinical settings and on the perspectives of minority Muslim populations in non-Muslim-majority countries as stated in Figure 2.



Topic/Outcome	Premarital Screening	Prenatal Diagnosis	Gene Editing	Community Attitudes	Policy/Regulation
Disease Prevention	12	10	2	4	6
Non-medical Enhancement	1	GAP	3	2	1
Ethical/Legal Analysis	5	6	8	3	7
Counseling Practices	7	5	GAP	6	2
Minority/Migrant Populations	2	1	GAP	2	1

Figure 2: Matrix of Research Topics and Study Attributes generated using *Consensus AI Tools*

CONCLUSION

Genetic testing in Islamic contexts offers important benefits for disease prevention but also raises ethical, legal, and social concerns that must be addressed in accordance with Islamic values. Continued collaboration among scholars, healthcare professionals, and policymakers is essential to ensure that these technologies are applied responsibly and respectfully within Muslim communities.

Future research should focus on how Muslim minorities in non-Muslim-majority countries experience genetic testing and counselling, the long-term social and psychological effects of mandatory premarital screening, and practical ways to apply Islamic bioethical principles in clinical and policy settings. These areas of inquiry, together with broader longitudinal and empirical research aimed at filling existing gaps in the literature, are vital for creating healthcare approaches that are both effective and culturally appropriate.

ACKNOWLEDGEMENT

The authors would like to express their sincere appreciation to the International Islamic University Malaysia (IIUM) for the continuous support throughout this study. We also extend our gratitude to all participants and collaborators whose contributions made this research possible.

REFERENCES

Abu-Shaheen, A., Dahan, D., Henaa, H., Nofal, A., Abdelmoety, D. A., Riaz, M., AlSheef, M., Almatary, A., & AlFayyad, I. (2022). Sickle cell disease in Gulf Cooperation Council countries: A systematic review. *Expert Review of Hematology, 15*(10), 893–909. https://doi.org/10.1080/17474086.2022.2132225

Alahmad, G., & Dierickx, K. (2018). Ethics of research biobanks: Islamic perspectives. Biopreservation and Biobanking, 16(3), 179–185. https://doi.org/10.1089/bio.2017.0067

Al-Aqeel, A. (2005). Ethical guidelines in genetics and genomics. An Islamic perspective. *Saudi Medical Journal, 26*(12), 1862–1870.

Alkuraya, F. S. (2014). Genetics and genomic medicine in Saudi Arabia. Molecular Genetics & Genomic Medicine, 2(5), 369–378. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4190871/

Amalia, A., & Hilmi, N. (2024). Etika pengobatan/rekayasa genetik dalam Islam sebagai implikasi untuk terapi gen dan teknologi DNA serta tantangan kontemporer. *Jurnal Medika Nusantara*. https://doi.org/10.59680/medika.v2i4.1479

Aqeel, A. (2007). Islamic ethical framework for research into and prevention of genetic diseases. *Nature Genetics, 39*, 1293–1298. https://doi.org/10.1038/ng.2007.14

Arijo, Z. (2020). U-4 The importance & significance of DNA evidence in Islamic Shari'ah. *Al-Aijaz Research Journal of Islamic Studies & Humanities*. https://doi.org/10.53575/u4.v4.02(20).63-80

Arijo, Z., & Khan, H. (2020). *Habibia Islamicus*. http://habibiaislamicus.com/index.php/hirj/article/view/148

Chin, A., Al-Balas, Q., Ahmad, M., Alsomali, N., & Ghaly, M. (2023). Islamic perspectives on polygenic testing and selection of IVF embryos (PGT-P) for optimal intelligence and other non–disease-related socially desirable traits. *Journal of Bioethical Inquiry, 21*, 441–448. https://doi.org/10.1007/s11673-023-10293-0

Ghaly, M. (2012). *Religio-ethical discussions on organ donation among Muslims in Europe: An example of transnational Islamic bioethics. Medicine, Health Care and Philosophy, 15*(2), 207–220. https://doi.org/10.1007/s11019-011-9352-x

Khan, M., & Konje, J. (2019). Ethical and religious dilemmas

of modern reproductive choices and the Islamic perspective. *European Journal of Obstetrics, Gynecology, and Reproductive Biology, 232*, 5–9. https://doi.org/10.1016/j.ejogrb.2018.10.052

Muhsin, S., Arab, S., & Chin, A. (2023). Islamic viewpoints on opportunistic sex selection of IVF embryos upon doing preimplantation genetic testing for preventing genetic diseases. *Asian Bioethics Review*, 1–10. https://doi.org/10.1007/s41649-023-00258-1

Novikova, R. (2019). Islam and genetics: Religious, ethical and legal issues. *Russian Journal of Philosophical Sciences, 23*, 565–585. https://doi.org/10.22363/2313-2337-2019-23-4-565-585

Padela, A., & Auda, J. (2020). *The moral status of organ donation and transplantation within Islamic law: The Fiqh Council of North America's position. Transplantation Direct, 6*(9), e598. https://doi.org/10.1097/txd.0000000000000980

Sachedina, A. (2020). Islamic ethics and the genome question. *Journal of Islamic Studies*. https://doi.org/10.1093/jis/etaa034

Shabana, A. (2017). Empowerment of women between law and science. *Hawwa, 15*, 193–218. https://doi.org/10.1163/15692086-12341327

Shabana, A. (2020). Islamic ethics and the legitimacy of scientific innovation. *Sociology of Islam, 8*, 265–289. https://doi.org/10.1163/22131418-00802006

Shabana, A. (2022). Between treatment and enhancement: Islamic discourses on the boundaries of human genetic modification. *Journal of Religious Ethics*. https://doi.org/10.1111/jore.12404

Shanti, H., Chouchane, L., Badii, R., Gallouzi, I., & Gasparini, P. (2015). Genetic testing and genomic analysis: A debate on ethical, social and legal issues in the Arab world with a focus on Qatar. *Journal of Translational Medicine, 13*. https://doi.org/10.1186/s12967-015-0720-9

Shoaib, A. (2024). Islamic perspectives on preconception, prenatal, and perinatal counseling. *Frontiers in Pediatrics, 12*. https://doi.org/10.3389/fped.2024.1373918

Sulaiman, K. (2018). Pre-marital test of genetic compatibility for effective child health: An Islamic perspective. *Alqalam Journal of Medical and Applied Sciences, 2*(1), 1–12. https://journal.utripoli.edu.ly/index.php/Alqalam/article/view/105

Tadmouri, G. O., Nair, P., Obeid, T., Al Ali, M. T., Al Khaja, N., & Hamamy, H. (2009). Consanguinity and reproductive health among Arabs. *Reproductive Health*, 6(1), 17. https://doi.org/10.1186/1742-4755-6-17

REVIEW ARTICLE

The Vicious Cycle of Menopause: A Narrative Review of the Interplay Between Urinary Incontinence, Mood Disturbances, and Sleep

Nur Ain Syuhada Mohd Zaini[1], Mohd Yusof Mohamad[1], Mohamed Arshad Mohamed Sideek[1], Ahmad Fahmi Harun Ismail[1,*]

[1]Department of Physical Rehabilitation Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia

Keywords:
menopause; urinary incontinence; depression; anxiety; midlife women.

ABSTRACT

Background: The menopausal transition is a period marked by the significant co-occurrence of urinary incontinence (UI) and mood disturbances (MD), such as depression and anxiety, which substantially impair quality of life. While often attributed separately to hormonal decline, the underlying mechanisms for their comorbidity are not fully explained by oestrogen deficiency alone. A growing body of evidence suggests a complex interplay, potentially mediated by other factors, with sleep disturbance emerging as a critical candidate. **Methods:** This narrative review synthesized literature identified through a comprehensive search of PubMed, Scopus, and PsycINFO databases. The search strategy combined keywords related to menopause, urinary incontinence, mood disorders, and sleep. We included observational studies, clinical trials, and relevant reviews published in English, focusing on peri- and postmenopausal women. The analysis was structured to evaluate the bidirectional relationships between UI, MD, and sleep, and to assess the evidence for sleep's role as a mediating factor. **Results:** The synthesis of evidence reveals strong, bidirectional relationships between all elements of the proposed "menopausal triad." UI, particularly through nocturia, directly fragments sleep architecture, leading to decreased sleep efficiency and reduced restorative slow-wave sleep. This sleep disruption, in turn, provokes neurobiological changes (e.g., amygdala hyperactivity, prefrontal cortex dysfunction) and HPA-axis dysregulation that predispose to and exacerbate MD. Concurrently, MD heightens the perception of UI severity through somatic awareness and catastrophizing, while also reducing adherence to self-management strategies. Critically, sleep disturbance functions as a central mediator, creating a self-perpetuating cycle where UI worsens sleep, which in turn worsens mood, ultimately increasing the burden and perception of UI. **Conclusion:** This review establishes that UI and MD in midlife women are intricately linked through the mediating pathway of sleep disturbance. The "menopausal triad" model provides a transformative framework for clinical practice, moving beyond a siloed approach. It mandates integrated assessment, where screening for all three components is essential, and promotes treatment strategies that target sleep improvement as a powerful leverage point to break the vicious cycle. This holistic approach is vital for allied health professionals to improve the overall quality of life for women navigating midlife.

INTRODUCTION

The menopausal transition is a pivotal life stage characterised by profound endocrinological changes that precipitate a wide array of symptoms, significantly impacting a woman's physical and psychological well-being (Monteleone et al., 2018). Among the most prevalent and distressing symptoms are urinary incontinence (UI) and mood disturbances (MD), including major depressive disorder and generalised anxiety disorder (Gibson et al., 2019). Epidemiological data consistently reveal that UI and MD co-occur at rates significantly higher than chance, suggesting a shared pathophysiology beyond common hormonal origins (Huang et al., 2020).

The conventional clinical narrative often attributes the parallel emergence of UI and MD to declining oestrogen levels. Oestrogen deficiency contributes to urogenital atrophy, decreased urethral closure pressure, and altered bladder function (Kim et al., 2021). It also simultaneously modulates central nervous system neurotransmitters like serotonin and norepinephrine, which are critical for mood regulation (Soares, 2017). However, this reductionist hormonal model is insufficient. Many women experience severe UI without MD, and vice versa, indicating the involvement of other critical, modifiable mediators (Avis et al., 2018).

This narrative review proposes that sleep disturbance is a crucial and underappreciated mediator that fuels a powerful bidirectional cycle between UI and MD. We synthesise evidence demonstrating that UI is a direct cause of sleep fragmentation (Perez-Lopez et al., 2021) and that poor sleep is, in itself, a robust and independent predictor for the onset and exacerbation of MD (Joffe et al., 2019;

* *Corresponding author.*
E-mail address: ahmadfahmi@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

Palagini et al., 2022). Conversely, MD can exacerbate the subjective burden of UI and perpetuate sleep problems through cognitive and behavioural pathways (Narrow et al., 2022). The primary objective of this review is to integrate findings from urology, psychiatry, and sleep medicine to build a compelling case for this "menopausal triad." By presenting this conceptual model, we aim to foster a holistic clinical approach and stimulate research into multimodal interventions that target the entire cycle.

FROM URINARY INCONTINENCE TO SLEEP DISRUPTION

The pathway from UI to poor sleep is both direct and multifactorial. Nocturia, or the need to wake at night to void, is a primary mechanism. Each awakening fragments the natural sleep cycle, significantly reducing time spent in deep, restorative (N3) sleep and rapid eye movement (REM) sleep, which are critical for cognitive function and emotional homeostasis (Tikkinen et al., 2010; Bliwise et al., 2021).

The burden of nocturia extends beyond frequency; its bother and the associated worry significantly correlate with poorer sleep quality independent of the number of voids (Perez-Lopez et al., 2021). Beyond the physiological trigger of nocturia, the constant fear of leakage can lead to a state of heightened vigilance and pre-emptive waking, even in the absence of a full bladder. This state of cognitive hyperarousal, characterized by anxiety about incontinence, is a core feature of insomnia and actively prevents sleep initiation and maintenance (Roth, 2020).

FROM SLEEP DISRUPTION TO MOOD DISTURBANCES

The neurobiological consequences of chronic sleep deprivation create a fertile ground for MD. Sleep is fundamental for effective emotional regulation; sleep loss is associated with increased amygdala reactivity (the brain's fear centre) and decreased functional connectivity with the prefrontal cortex, which governs executive function and emotional control (Goldstein & Walker, 2014). This neural dysregulation manifests behaviourally as increased irritability, emotional lability, and a pronounced negative cognitive bias, all of which are core features of depression and anxiety disorders (Palagini et al., 2022). Furthermore, sleep disruption reliably dysregulates the hypothalamic-pituitary-adrenal (HPA) axis, leading to elevated cortisol levels and a blunted circadian rhythm, which are established physiological correlates of depressive disorders (Kalmbach et al., 2018). In menopausal women, sleep disruption arising from vasomotor instability may act synergistically with UI-related nocturnal awakenings, producing an amplified adverse cascade on mood and affective stability (Joffe et al., 2019).

THE INFLUENCE OF MOOD ON URINARY INCONTINENCE AND SLEEP

It is critical to recognize that the relationship within the triad is not unidirectional. Mood disorders can significantly influence the perception, severity, and management of UI and are, themselves, primary drivers of sleep disturbance. Depression and anxiety are strongly associated with heightened somatic awareness and catastrophizing, cognitive processes that can amplify the perceived severity and distress associated with UI symptoms (Narrow et al., 2022). Behaviourally, low mood, anhedonia, and fatigue can profoundly reduce motivation and adherence to effective self-management strategies, such as consistent pelvic floor muscle training (Cody et al., 2012). From a physiological perspective, chronic anxiety can increase overall muscular tension, including dysfunctional tension in the pelvic floor muscles, potentially worsening symptoms of specific UI subtypes (Lone et al., 2020). Finally, MD are established primary causes of sleep-onset and maintenance insomnia, thereby directly entrenching the cycle of poor sleep and ensuring its self-perpetuating nature (Roth, 2020).

METHODS

Review Design and Aim

This study employed a narrative review methodology to synthesize and critically appraise the existing literature on the interplay between urinary incontinence (UI), mood disturbances (MD), and sleep in midlife women. The primary aim was to develop a novel conceptual model—the "menopausal triad"—that positions sleep as a central mediator. Unlike a systematic review, which aims for a comprehensive, protocol-driven aggregation of all empirical evidence on a specific, narrow question, a narrative review allows for a broader, more exploratory synthesis of a complex and interdisciplinary topic, enabling the identification of themes, gaps, and the development of new theoretical frameworks (Ferrari, 2015; Green et al., 2006).

Search Strategy and Data Sources

To ensure a rigorous and comprehensive literature search, we queried three major electronic databases: PubMed/MEDLINE, Scopus, and PsycINFO. The search strategy was designed to capture the core concepts of the population and the key variables of interest. A combination of Medical Subject Headings (MeSH) terms

and free-text keywords was used, including: ("menopause" OR "perimenopause" OR "postmenopause" OR "midlife women") AND ("urinary incontinence" OR "nocturia" OR "overactive bladder") AND ("mood disorders" OR "depression" OR "anxiety" OR "affective symptoms") AND ("sleep" OR "sleep wake disorders" OR "sleep quality" OR "insomnia"). The search was limited to articles published in English from January 2000 to December 2023. Reference lists of key articles were also hand-searched to identify additional relevant publications.

Eligibility Criteria

Studies were selected based on the following criteria:

- Population: Peri- or postmenopausal women (typically aged 40-65).
- Concepts: Studies had to investigate a relationship between at least two of the three core elements of the triad (UI, MD, and sleep).
- Study Types: Original research articles (cross-sectional, case-control, longitudinal cohort studies, and clinical trials), systematic reviews, and meta-analyses were included to provide a comprehensive overview of evidence and established consensus.

The exclusion criteria were: (1) studies focusing on populations with specific comorbid neurological diseases (e.g., Parkinson's, multiple sclerosis) or active cancer treatment, as these conditions could independently and severely affect all three triad components; (2) studies not published in English; and (3) editorials, opinion pieces, and single case reports.

Selection Process and Data Synthesis

The literature selection and synthesis process followed the phases outlined in the established guidance for narrative reviews by Green et al. (2006) and Ferrari (2015), as depicted in Figure 1. The process began with the initial database search, which yielded a large volume of records. After removing duplicates, the titles and abstracts of the remaining records were screened for relevance against the eligibility criteria. The full texts of potentially relevant articles were then retrieved and thoroughly assessed to make the final inclusion decision.

Given the narrative synthesis approach, data was not extracted for meta-analysis. Instead, the included literature was analysed thematically to map the key pathways and interactions between UI, MD, and sleep. The synthesis was structured to first examine the direct relationships between each pair of variables (UI-Sleep, Sleep-MD, MD-UI) before integrating these findings to build the case for sleep's role as a central mediator in the proposed triad model. The quality of included studies was considered during the interpretation and discussion of the findings, with greater weight given to findings from longitudinal studies and high-quality systematic reviews.

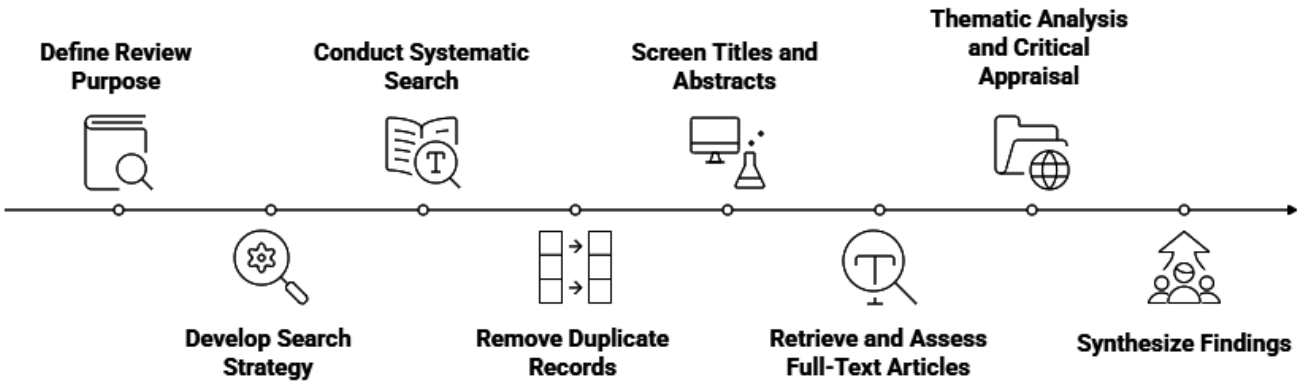


Figure 1: The literature selection and synthesis process.

SYNTHESIZING THE EVIDENCE: SLEEP AS THE CENTRAL MEDIATOR

When these individual pathways are considered together, a compelling and clinically useful model emerges. We propose that sleep is not merely a secondary symptom but a central mechanistic link that binds UI and MD in a self-perpetuating cycle. Consider a typical sequence: a woman experiences nocturia due to UI, leading to fragmented and non-restorative sleep. This sleep deprivation impairs her prefrontal cortex function and emotional resilience the following day, increasing her susceptibility to low mood, irritability, and anxiety (Baglioni et al., 2011). This

worsened mood, in turn, increases catastrophic thinking regarding her urinary symptoms and reduces her energy for pelvic floor exercises. These effects potentially worsen UI severity, leading to another night of poor sleep and further mood deterioration. This vicious cycle, illustrated in Table 1, powerfully demonstrates how sleep disturbance can act as the engine that drives the comorbidity of UI and MD, moving beyond a simple correlation to a dynamic pathological process.

Table 1: *The Vicious Cycle of the Menopausal Triad: Exemplar Pathways*

Stage	Symptom Manifestation	Underlying Mechanism	Consequence
1	Increased urinary incontinence, particularly nocturia.	Urogenital atrophy, decreased bladder capacity, detrusor overactivity.	Frequent nighttime awakenings, sleep fragmentation, reduced slow-wave sleep.
2	Sleep disruption and deprivation.	Reduced slow-wave and REM sleep; HPA axis dysregulation; increased inflammatory markers.	Impaired emotional regulation, increased negative affect, fatigue, hyperalgesia.
3	Worsening mood (depression, anxiety).	Amygdala hyperactivity, prefrontal cortex dysfunction, noradrenergic/serotonergic imbalance.	Heightened somatic awareness, catastrophizing about UI, reduced self-efficacy, social withdrawal.
4	Perceived worsening of UI burden.	Behavioural withdrawal from self-care (e.g., pelvic floor exercises); anxiety-induced pelvic floor tension.	Potential physiological worsening of UI; increased focus on symptoms.
5	Cycle Repeats	Return to Stage 1 with heightened severity across all domains.	Consolidation of the vicious cycle leads to a significant decline in quality of life.

CLINICAL IMPLICATIONS AND A CALL FOR INTEGRATED CARE

The "menopausal triad" model has profound implications for clinical practice, mandating a shift away from symptom-specific silos toward a more holistic, patient-centred approach. A fundamental implication is the necessity for systematic, holistic assessment. A woman presenting with one element of the triad should be automatically screened for the other two. For example, a gynecologist or primary care physician managing a patient's UI should use brief, validated tools like the Patient Health Questionnaire-9 (PHQ-9) for depression and the Pittsburgh Sleep Quality Index (PSQI) for sleep (Kroenke et al., 2001; Buysse et al., 1989). Similarly, a psychiatrist or therapist treating a perimenopausal woman for new-onset anxiety should proactively inquire about UI and sleep quality. This cross-disciplinary screening is the first step in identifying the presence of the full cycle.

Following assessment, treatment plans must be integrated and designed to target multiple components of the triad simultaneously for maximum effect. Effectively managing UI through evidence-based interventions such as pelvic floor physical therapy, pharmacotherapy (e.g., anticholinergics, beta-3 agonists), or neuromodulation can directly reduce nocturia, thereby directly improving sleep quality and indirectly alleviating mood symptoms (Kim et al., 2021). Conversely, treating MD with cognitive-behavioural therapy (CBT) or appropriate medication can improve sleep quality and enhance the patient's psychological capacity and motivation to manage her UI (Soares, 2017) proactively.

Perhaps most promisingly, directly addressing sleep disturbance with targeted interventions like Cognitive Behavioural Therapy for Insomnia (CBT-I) could serve as a powerful primary intervention to disrupt the entire cycle, potentially leading to concurrent improvements in both

mood and the perceived burden of UI (Kalmbach et al., 2018). Table 2 summarizes these multidimensional strategies, underscoring the need for future clinical frameworks that operationalize integrated care for menopausal health.

Table 2: *Clinical Considerations for Managing the Menopausal Triad*

Clinical Element	Key Considerations	Suggested Tools & Interventions
Assessment & Screening	Screen for all three components regardless of presenting complaint. Implement cross-disciplinary screening protocols.	**UI:** ICIQ-UI SF, 3-day bladder diary. **Mood:** PHQ-9, GAD-7. **Sleep:** PSQI, ISI.
Integrated Treatment Planning	Develop a collaborative, multi-component treatment plan. Treating one component can positively impact the others.	**UI:** Pelvic floor physical therapy, anticholinergics/beta-3 agonists. **Mood:** CBT, SSRIs/SNRIs. **Sleep:** Sleep hygiene, CBT-I.
Patient Education & Empowerment	Validate the patient's experience by explaining the interconnected nature of her symptoms. Frame treatment as breaking a cycle.	Use the triad model to illustrate the vicious cycle. Set realistic goals for incremental improvement across domains. Emphasize that these symptoms are a common, treatable part of midlife health.

Note: ICIQ-UI SF = International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form; PHQ-9 = Patient Health Questionnaire-9; GAD-7 = Generalised Anxiety Disorder-7; PSQI = Pittsburgh Sleep Quality Index; ISI = Insomnia Severity Index; CBT = Cognitive Behavioural Therapy; CBT-I = Cognitive Behavioural Therapy for Insomnia; SSRIs/SNRIs = Selective Serotonin Reuptake Inhibitors/Serotonin-Norepinephrine Reuptake Inhibitors.

CONCLUSION AND FUTURE DIRECTIONS

In conclusion, the interplay between urinary incontinence and mood disturbances in midlife women is best understood through a triad model that centrally incorporates sleep disturbance. Viewing these conditions as interconnected parts of a pathological cycle, rather than as isolated sequelae of oestrogen decline, provides a more accurate, comprehensive, and compassionate framework for patient care. Recognizing sleep as a key mediating factor offers a tangible and modifiable target for clinical intervention, providing a new avenue for improving outcomes.

Future research should prioritize longitudinal studies to confirm the temporal relationships and causal pathways implied in this model. Employing sophisticated statistical methods, such as cross-lagged panel analysis and formal mediation analysis, will be crucial to quantify the relative contribution of sleep within the triad. Furthermore, interventional trials are critically needed to test the efficacy of combined treatment approaches. For instance, a randomized controlled trial comparing the outcomes of pelvic floor therapy combined with CBT-I against standard, single-symptom care would provide robust evidence for the integrated model. Investigating the role of novel therapies, such as mindfulness-based stress reduction, which may simultaneously target mood, sleep, and the cognitive perception of UI, is another promising direction. By embracing the complexity of the menopausal experience, clinicians and researchers can forge person-centred strategies that not only break this vicious cycle but also promote enduring well-being in midlife and beyond.

ACKNOWLEDGEMENT

This research was not funded by any grant, and there are no known conflicts of interest between the authors.

REFERENCES

Avis, N. E., Crawford, S. L., & Greendale, G. (2018). Duration of menopausal vasomotor symptoms over the menopause transition. *JAMA Internal Medicine, 178*(8), 1121–1122.

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., Lombardo, C., & Riemann, D. (2011). Insomnia as a predictor of depression: A

meta-analytic evaluation of longitudinal epidemiological studies. *Journal of Affective Disorders, 135*(1-3), 10–19.

Bliwise, D. L., Wagg, A., & Sand, P. K. (2021). Nocturia: A highly prevalent disorder with multifaceted consequences. *Urology, 157*, 1-9.

Buysse, D. J., Reynolds, C. F., 3rd, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213.

Cody, J. D., Jacobs, M. L., Richardson, K., Moehrer, B., & Hextall, A. (2012). Oestrogen therapy for urinary incontinence in post-menopausal women. *Cochrane Database of Systematic Reviews, 10*, CD001405.

Ferrari, R. (2015). Writing narrative style literature reviews. *Medical Writing, 24*(4), 230-235. https://doi.org/10.1179/2047480615Z.000000000329

Gibson, C. J., Li, Y., Bertenthal, D., Huang, A. J., & Seal, K. H. (2019). The association between anxiety, depression, and urinary incontinence among community-dwelling women. *American Journal of Obstetrics and Gynecology, 221*(4), 347.e1-347.e10.

Goldstein, A. N., & Walker, M. P. (2014). The role of sleep in emotional brain function. *Annual Review of Clinical Psychology, 10*, 679–708.

Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. *Journal of Chiropractic Medicine, 5*(3), 101–117. https://doi.org/10.1016/S0899-3467(07)60142-6

Huang, A. J., Gregorich, S. E., Kuppermann, M., Nakagawa, S., Van Den Eeden, S. K., Brown, J. S., Richter, H. E., Walter, L. C., & Thom, D. H. (2020). Day-to-day impact of urinary incontinence and its association with depression in middle-aged and older women. *Maturitas, 141*, 47–52.

Joffe, H., Crawford, S., Freeman, M., White, D., & Soares, C. N. (2019). Impact of sleep disturbance on health and functioning in midlife women: A narrative review. *Menopause, 26*(9), 1055-1061.

Kalmbach, D. A., Anderson, J. R., & Drake, C. L. (2018). The impact of stress on sleep: Pathogenic sleep reactivity as a vulnerability to insomnia and circadian disorders. *Journal of Sleep Research, 27*(6), e12710.

Kim, D. K., Lee, J. Y., & Ryu, D. S. (2021). The effect of menopause on the lower urinary tract function: A nationwide epidemiologic study. *Journal of Menopausal Medicine, 27*(2), 78-85.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613.

Lone, F. W., Rashid, P., & Sultan, A. H. (2020). The relationship between pelvic floor muscle strength and sexual function in women with stress urinary incontinence. *Journal of Sexual Medicine, 17*(6), 1107-1115.

Monteleone, P., Mascagni, G., Giannini, A., Genazzani, A. R., & Simoncini, T. (2018). Symptoms of menopause — Global prevalence, physiology and implications. *Nature Reviews Endocrinology, 14*(4), 199–215.

Narrow, W. E., Clarke, D. E., & Kuramoto, S. J. (2022). DSM-5 field trials in the United States and Canada, Part III: Development and reliability testing of a cross-cutting symptom assessment for DSM-5. *American Journal of Psychiatry, 179*(1), 45-55.

Palagini, L., Hertenstein, E., Riemann, D., & Nissen, C. (2022). Sleep, insomnia and mental health. *Journal of Sleep Research, 31*(4), e13628.

Perez-Lopez, F. R., Ornat, L., Lopez-Baena, M. T., & Perez-Roncero, G. R. (2021). Association of nocturia and sleep problems with depressive symptoms in perimenopausal women. *Climacteric, 24*(4), 369-375.

Roth, T. (2020). Insomnia: Definition, prevalence, etiology, and consequences. *Journal of Clinical Sleep Medicine, 3*(5 Suppl), S7–S10.

Soares, C. N. (2017). Depression and menopause: Current knowledge and clinical recommendations for a critical window. *Psychiatric Clinics of North America, 40*(2), 239–254.

Tikkinen, K. A., Johnson, T. M., 2nd, Tammela, T. L., Sintonen, H., Haukka, J., Huhtala, H., & Auvinen, A. (2010). Nocturia frequency, bother, and quality of life: How often is too often? A population-based study in Finland. *European Urology, 57*(3), 488–496.

ORIGINAL ARTICLE

University Students and Excessive Screen Time: Links to Sleep Disturbance and Acute Health Issues

Najiyya Abdullah Kamil[1], Norfaezah Ahmad[2], Norhidayah Ahmad[1,*]

[1] Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia.
[2] Department of Prosthodontics, Kulliyyah of Dentistry, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
screen time; sleep quality; health effect; university students; smartphone addiction

ABSTRACT

Background: University students often spend a significant amount of time on screens due to academic demands, entertainment and social interactions. However, excessive screen time may negatively affect sleep quality and contribute to deterioration in health. This study aimed to determine the association between sleep disturbance and acute health effects associated with prolonged screen time among students at a higher education institution. **Methods:** A cross-sectional study was conducted among undergraduate students at a public university in Kuantan, Pahang. A self-administered questionnaire was distributed to eligible participants who owned personal screen-based devices and had sufficient English proficiency to understand the questionnaire. Students who were on study leave and had been diagnosed with insomnia and pre-existing eye disorders that are known not to be caused by excessive screen usage were excluded. Screen device addiction was measured using the Manolis/Roberts Cell-Phone Addiction Scale (MRCPAS), while sleep behaviour was assessed using the Pittsburgh Sleep Quality Index (PSQI). **Results:** With 155 participants, the findings indicate that the majority preferred using smartphones primarily for educational and entertainment purposes, with an average daily screen time of 7.32 hours. The mean addiction score to the screen device was 12.81 out of 20, indicating a high level of screen device dependency. Despite this, the majority (97.4%) reported good sleep quality. Participants reported several acute health effects due to prolonged screen time, including eye strain, musculoskeletal discomfort (e.g. back pain, neck pain), headaches, and a decline in physical fitness. Chi-square analysis revealed significant associations between sleep disturbance and dry eyes (p=0.022), blurred vision (p=0.023), burning eyes sensation (p=0.048), double vision (p=0.014), difficulty focusing for near vision (p=0.012), feeling that eyesight is worsening (p=0.029), and back pain (p=0.028). **Conclusion:** Our findings indicate that sleep disturbance had significant associations with several acute health effects caused by prolonged screen time. The results obtained in this study may not accurately reflect the general university student population, as the sample size was limited; however, they may serve as a helpful starting point for larger epidemiological studies.

INTRODUCTION

University students are often committed to having high screen time in their daily lives, as they must complete assignments, attend group discussions, and sometimes participate in online classes. Student's screen time should be controlled because excessive screen use can harm their physical health and well-being. Excessive screen time was found to be prone to developing many adverse health effects, such as headaches, eye problems, musculoskeletal disorders, cardiovascular disease, mental health issues and sleep disturbance.

The World Health Organisation [WHO] (2019) describes screen time as the "*time spent passively watching screen-based entertainment (TV, computer, mobile devices) and does not include active screen-based games where physical activity or movement is required*". This statement is also commonly referred to as sedentary screen time, in which screen time activities are performed in a resting position, such as sitting or lying down (Bull et al., 2020). Ross et al. (2020) recommend that adults limit sedentary activities to 8 hours or fewer per day, increase physical activity, and adopt a healthy lifestyle that includes regular exercise.

Focusing on screen time, screen-based activities differ across individuals because screens can perform various functions simultaneously. In daily life, electronic devices are primarily used to facilitate tasks such as socialising, searching for knowledge, and entertainment (Kaye et al., 2020). With the readily available internet connection in recent years, compared to the past, Malaysians have spent most of their time on screens browsing or surfing the internet. According to The New Straits Times (2023),

* *Norhidayah Ahmad.*
E-mail address: hidayahahmad@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

Malaysia ranked third among the ten leading nations with the highest internet usage frequency, averaging eight hours and six minutes daily. It has been stated that almost all citizens of Malaysia are active online users, with an estimated 96.8 per cent, and this is one of the elements that leads this country to occupy such a high rank (New Straits Times, 2023).

Impact of Excessive Screen Time on Sleep Quality

Excessive screen time may disrupt the sleep cycle and reduce sleep quality. Research conducted by Hale et al. (2018) confirms a relationship between excessive screen time and poor sleep quality, increased bedtime delay, and heightened daytime exhaustion. A study conducted by Cain & Gradisar (2010, as cited in Perrault et al., 2019) hypothesised three possibilities could be involved in how screen time affects sleep: (i) screen activities that usually last long will mess up the person's sleep schedule; (ii) raising feelings of emotion at night that influence sleep time; and (iii) screen light that prevents the synthesis of melatonin, a hormone known to promote sleep in the human body. However, subjection to light from 12:00 a.m. to 4:00 a.m. may restrict the release of the melatonin hormone, where this early light detection will accelerate the circadian rhythm, which causes the production of melatonin to reach its highest point ahead of time, yet light exposure in the evening delays the stage of melatonin secretion (Poza et al., 2022).

Screen activities, such as playing online games, can be extremely exciting and require intense concentration to succeed (Hale et al., 2018). Thus, this can result in delayed bedtimes for students. Being active on a screen, especially at night, can shorten the sleep cycle. When students do not get enough sleep, it can disrupt their daily lives and lead to poor academic performance. Insufficient sleep is known to worsen a person's medical conditions, impair cognitive ability, elevate stress and depression, and also weaken the body's immunity, which raises the possibility of getting injured (Alshoaibi et al., 2023).

Acute Health Effects of Excessive Screen Time on the Human Body

A study by Kaur et al. (2022) describes digital eye strain as eye or vision concerns arising from prolonged use of screen devices. This digital eye strain can be divided into two categories, which are ocular symptoms such as dry eyes, itching, tearing, tired eyes, blurred vision, burning sensation, redness and double vision and the associated non-ocular symptoms like stiff neck, general fatigue, headache, and back pain (Kaur et al., 2022; Agarwal et al., 2022). In addition to excessive screen usage, digital eye strain, also known as computer vision syndrome, has been linked to several issues, including poor sitting posture, an inappropriate distance between the eyes and the screen, and an incorrect angle of view for the screen user (Lema & Anbesu, 2022). The American Optometric Association noted that two hours of non-stop screen time is enough to cause digital eye strain (Kaur et al., 2022). A study conducted by AlQarni et al. (2023) found that 68.5% of university students experienced digital eye strain due to back-to-back online classes during the COVID-19 pandemic, leading to prolonged screen use. In addition, a study by Ccami-Bernal et al. (2023) found that the prevalence of digital eye strain was approximately 69.0%, with female university students affected more than male students.

Excessive screen time also contributes to a high prevalence of headaches among university students, including medical students (Montagni et al., 2016). In addition to one of the symptoms of digital eye strain, headaches are said to be caused by an improper distance between the user and the screen, and may also result from back muscle pain experienced by individuals who use screens excessively (Lucchetti, 2023; Shah & Hameed, 2023). Moreover, the prolonged screen time has been linked to musculoskeletal pain, particularly neck and back pain (Khan et al., 2021). As such, this study aims to assess screen behaviour and device addiction among university students at a public university and to determine the prevalence of sleep disturbances and related acute health effects associated with excessive screen time.

MATERIALS AND METHODS

Study Design and Participants

This cross-sectional study was approved by the IIUM Research Ethics Committee (ID: IREC 2024-KAHS/DBMS8). This study utilised convenience sampling to recruit undergraduate students from a public university in Kuantan, Pahang. Students were eligible to participate in this study if they owned any screen devices, such as smartphones, laptops, or tablets, and could read and understand English to complete the questionnaire. Students diagnosed with insomnia or eye problems known not to be caused by excessive screen use were excluded from the study.

Data Collection

Data collection was conducted from April to June 2024 using a self-administered questionnaire distributed via WhatsApp and Telegram to undergraduate students from various medical and health sciences programmes of a

public university. Participation was voluntary, and informed consent was obtained before completing the survey. All responses were anonymous and treated confidentially, and participants were free to withdraw at any time.

Questionnaire Design, Validity and Reliability

The questionnaire consisted of four sections, which are (i) socio-demographic characteristics; (ii) screen behaviour and addiction; (iii) sleep behaviour and (iv) acute health effects of excessive screen time. Screen behaviour and addiction were measured using the Manolis/Roberts Cell-Phone Addiction Scale (MRCPAS), while sleep behaviour was assessed using the Pittsburgh Sleep Quality Index (PSQI). In addition, several questions have been adapted from previous studies that have a similar scope to the aim of this study (Dwajani, 2021; Bharadwaj, 2021; Roberts et al., 2014; Buysse et al., 1989; Mohd Kamaruzihan & Soe, 2023; Agarwal et al., 2022; Seguí et al., 2015). Two experts in the field validated the questionnaire and underwent a pilot study to assess its reliability (SPSS, $\alpha = 0.721$).

Data Analysis

The data were analysed using SPSS. Descriptive statistics were used to summarise the frequency and percentage of screen time duration, types of screen devices, and purposes of screen time among participants. A chi-square test was applied to determine the association between categorical variables. Screen addiction was measured using questions adapted from the Manolis/Roberts Cell-Phone Addiction Scale (MRCPAS). Each respondent's score was obtained by summing the scores of four items, rated from 1 (Strongly Disagree) to 5 (Strongly Agree), with higher scores indicating greater screen addiction (maximum score of 20). Sleep behaviour was assessed using the Pittsburgh Sleep Quality Index (PSQI), which consists of seven components: sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbances, sleep medication, and daytime dysfunction. Each component is scored from 0 to 3, and the total score (0–21) indicates overall sleep quality. Based on the validated PSQI cut-off, a global score greater than five indicates poor sleep quality, while scores of 5 or below indicate good sleep quality (Insana et al., 2013; Park, 2020).

RESULTS

Response Rate and Socio-demographic Characteristics

The response rate obtained in this study was 43.0% (n=155) of the targeted sample size, which closely aligns with the weighted mean response rate of 44.1% (95% CI: 42.3%–46.0%) typically observed in online surveys within education-related fields as reported by Wu et al. (2022). The sample size was also considered sufficient based on a power calculation targeting 80% power to detect medium effect sizes at a significance level of 0.05. This sample size is also comparable to similar studies investigating screen time, sleep disturbances, and related health effects among university students (Mohd Kamaruzihan & Soe, 2023; Meghji et al., 2025).

The mean age was 22 years, and the average BMI was 22.36, indicating normal body weight (61.3%). Most participants were female (85.2%), reflecting the university's higher proportion of female students. About 41.3% were from the M40 income group (RM5,250–RM11,819 per month for the parents' household income). More than half (52.3%) were scholarship recipients, and almost all (97%) stayed in university accommodation. Nearly half (45%) reported having a sedentary lifestyle with little or no exercise per week.

Screen Behaviour and Addiction

All participants reported owning their own smartphones, and 92.3% own a laptop, while 78.7% own a tablet or iPad. University students used screen devices for a variety of purposes, with educational purposes being the most common (99.4%), followed by social media (97.4%) and entertainment purposes (96.1%). On average, daily screen time was 7.32 hours on smartphones, 4.29 hours on tablets, and 3.19 hours on laptops. The average recreational screen time among the participants was 4.54 hours.

As assessed by the screen addiction level through Manolis/Roberts Cell-Phone Addiction Scale (MRCPAS), the mean for the total score of screen addiction was 12.81, which ranged from 4 to 20, indicating that the average addiction level exceeded the midpoint value. Table 1 presents the participant responses to the screen addiction items. Most of the respondents disagreed (32.9%) or remained neutral (29.7%) regarding agitation when their smartphone or screen device was not in sight. Similarly, a moderate proportion reported feeling nervous when the device's battery was nearly depleted. In contrast, a majority of participants agreed (34.8%) or strongly agreed (24.5%) that they spent more time than intended on their devices and that their screen time had increased progressively.

Table 1: Screen Addiction Scores of study participants (Based on MRCPAS)

Questions	Response	N (%)
I feel agitated or uneasy when my smartphone/device is not within sight.	Strongly Disagree	20 (12.9)
	Disagree	**51 (32.9)**
	Neutral	**46 (29.7)**
	Agree	28 (18.1)
	Strongly Agree	10 (6.5)
I feel nervous or anxious when my smartphone/device's battery is running low.	Strongly Disagree	20 (12.9)
	Disagree	**43 (27.7)**
	Neutral	**43 (27.7)**
	Agree	37 (23.9)
	Strongly Agree	12 (7.7)
I often spend more time on my smartphone/device than I should.	Strongly Disagree	10 (6.5)
	Disagree	10 (6.5)
	Neutral	43 (27.7)
	Agree	**54 (34.8)**
	Strongly Agree	**38 (24.5)**
I have noticed that I am spending increasing amounts of time on my smartphone/device.	Strongly Disagree	10 (6.5)
	Disagree	16 (10.3)
	Neutral	**42 (27.1)**
	Agree	**47 (30.3)**
	Strongly Agree	40 (25.8)

Sleep Behaviour

In the assessment of sleep behaviour using the Pittsburgh Sleep Quality Index (PSQI), 97.4% of participants had good sleep quality, while only 2.6% had poor sleep quality. As illustrated in Figure 1, the PSQI questionnaire comprises seven components with the following mean scores: (i) 1.18 for quality, (ii) 1.23 for latency, (iii) 1.88 for duration, (iv) 0.43 for efficiency, (v) 1.15 for disturbance, (vi) 0.07 for medication, and (vii) 1.04 for daytime dysfunction. Each component yields a score of 0 to 3, and their sum yields a global PSQI score of 0 to 21. The present study reported an average global PSQI score of 2.48, indicating good overall sleep quality. A higher PSQI score reflects poorer sleep quality, and a global PSQI score greater than five is classified as poor sleep (Insana et al., 2013; Park, 2020).

Participants were also asked about their device usage before bedtime. Most of them reported frequently bringing the devices to bed and using them before sleep, primarily for browsing social media (82%) and texting or messaging (70%). In comparison, only a quarter of participants play games before sleep (25%). Additionally, the majority of students (94.2%) used their screen devices as alarms to wake up in the morning.

Acute Symptoms After Prolonged Screen Time

Figure 2 illustrates the acute health effects experienced by participants following prolonged screen use. More than half of the participants reported experiencing eye strain (52.3%), back pain (54.85%), neck pain (57.4%), headache (52.9%), and reduced physical strength (52.9%) after extended periods of screen exposure. The remaining participants either did not experience these symptoms or were not sure about them.

Other acute symptoms, such as excessive blinking, blurred vision, burning eyes sensation, double vision, eye irritation, tearing eyes, difficulty focusing for near vision, and feeling that eyesight is worsening, were only experienced by a small group of participants, ranging between 11% to 38% of the total study population. Meanwhile, the distribution of participant's opinions on "dry eyes" and "increased sensitivity to light" was almost the same between agree and disagree.

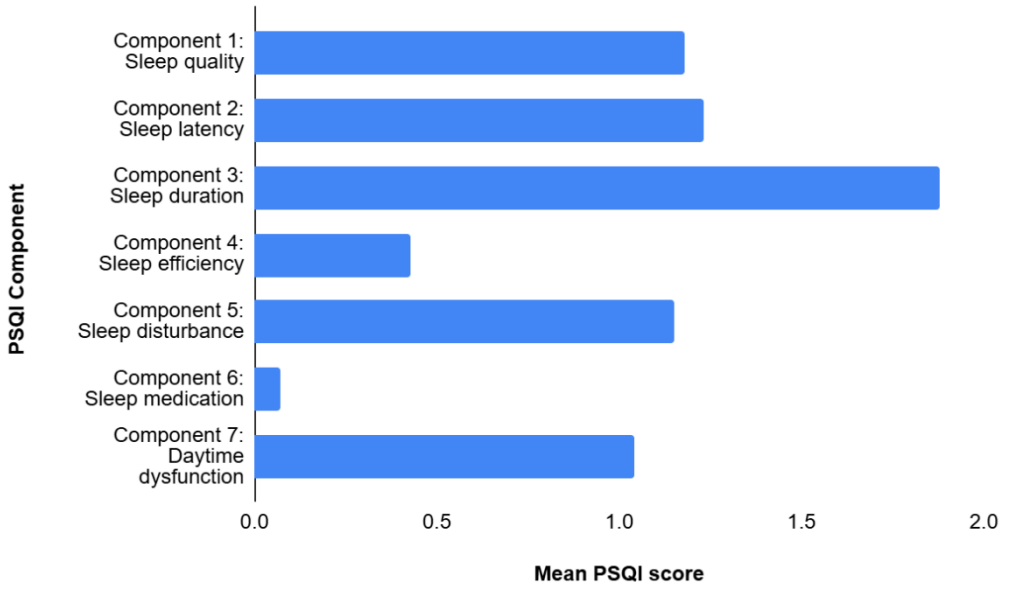


Figure 1: Mean PSQI Component Scores of Study Participants

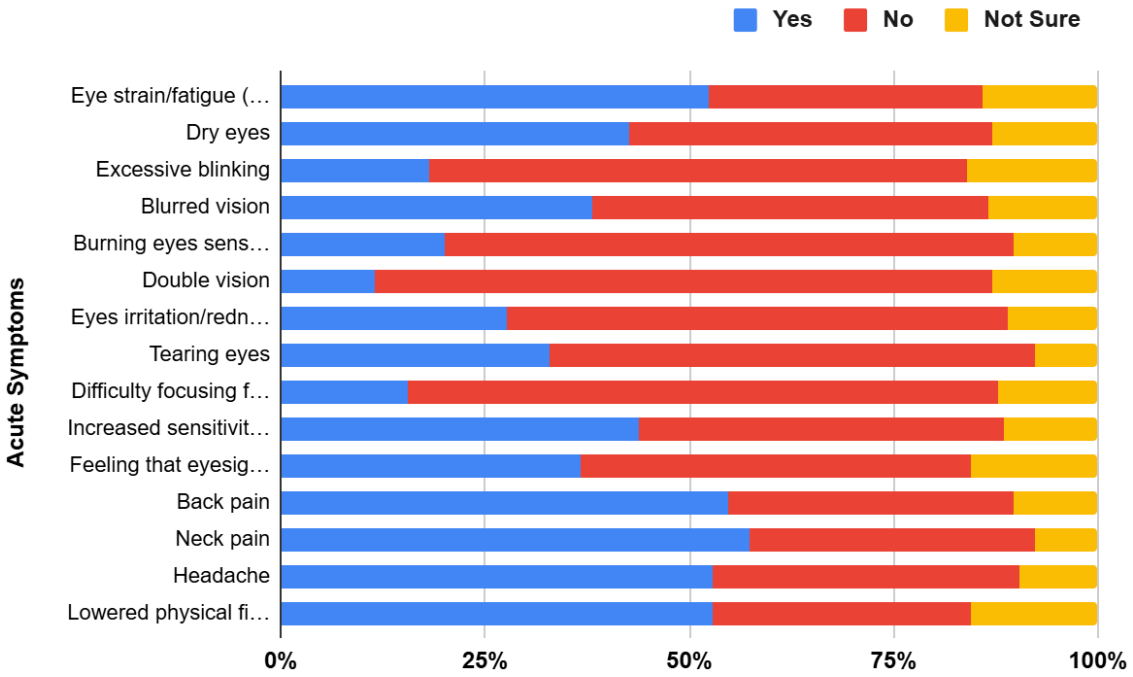


Figure 2: Acute Symptoms Reported by Study Participants Following Prolonged Screen Exposure

Association Between Sleep Disturbance and Acute Health Effects

The association between sleep disturbance and acute health effects following prolonged screen exposure has been assessed statistically. The results show significant associations with symptoms of dry eyes ($p = 0.022$), blurred vision ($p = 0.023$), burning eyes sensation ($p = 0.048$), double vision ($p = 0.014$), difficulty focusing for near vision ($p = 0.012$), feeling of worsening eyesight ($p = 0.029$) and back pain ($p = 0.028$) (Chi-square Test; n=155).

DISCUSSION

The present study reported that participants owned various devices, including smartphones, laptops, tablets, and iPads, similar to Mohd Kamaruzihan and Soe's (2023) study, which found that all university students at a public university owned these mobile gadgets. Nowadays, it has become a necessity for university students to have at least a smartphone for communication, online banking or digital payments. In addition, they use these devices for a variety of purposes, including education (online classes or discussions), social media, and entertainment. A study by

Ataş and H.Çelik (2019) reported that more than 80% of university students used smartphones for online searches, social media, and communication, including messaging and calling. The present study found that the mean screen time across all devices exceeded 3 hours (smartphones: 7.32 hours, laptops: 3.19 hours, and tablets/iPads: 4.29 hours), which was considered high and associated with poor health outcomes (Biaani et al., 2020).

Due to the prolonged screen time among the study population, this study hypothesised that the participants would experience sleep disturbances. However, as assessed by the PSQI, the results contradicted this assumption, with only 2.6% (n=4 out of 155) of participants reporting poor sleep quality (PSQI score of 2.48). These findings also contrast with the earlier study by Mohd Kamaruzihan and Soe (2023), who reported that 56% (n=56 out of 100) of pharmacy students from the same population experienced poor sleep (PSQI score of 6.7). The lower proportion of students experiencing sleep disturbances in the present study may be attributed to factors such as the data-collection period and the sample's demographic characteristics. As data collection occurred during the non-exam period, students were likely experiencing lower academic stress, which may have contributed to better sleep quality (Khan et al., 2025). In addition, the present study included students from various health-related programmes and study years. In contrast, the earlier study by Kamaruzihan and Soe (2023) focused solely on pharmacy students, who generally face heavier academic workloads, credit units and practical hours (Hanapi et al., 2021). This difference may partly explain the higher percentage of poor sleepers reported in their study.

The majority of participants in the present study report experiencing eye strain, back pain, neck pain, headaches, and reduced physical strength after prolonged screen time. Previous studies support these findings, including a study by Yıldız et al. (2023), which found that 58.8% of participants experienced reduced physical activity, 85% had low back pain, and 79.4% had neck pain as the aftereffects of prolonged screen time. In addition, 52.8% of participants reported eye strain and 31.3% reported headache after prolonged screen use in a study by Agarwal et al. (2022).

Association Between Sleep Disturbance Due to Prolonged Screen Time and Acute Health Effects

The present study found several statistically significant associations between sleep disturbance due to prolonged screen time and acute health symptoms, including dry eyes, blurred vision, burning eye sensations, double vision, difficulty focusing for near vision, worsening eyesight, and back pain (p-value < 0.05, Chi-square Test, n = 155). This finding was consistent with the previous study by Albalawi et al. (2023), which found that college students with more than 6 hours of screen time had severe sleep-wake disturbances, and that almost half of the participants experienced mild to moderate dry eye symptoms. They postulated that female sex hormones, particularly a decrease in estrogen, might play an essential role in the association with dry eye syndrome. As the majority of the participants in the present study were female (85.2%), this may have been an important confounding variable that affected the results.

A study by Chu et al. (2023) also found that students with high smartphone usage of more than 4 hours per day were more likely to experience digital eye syndrome, including blurred vision. In the present study, the mean duration of smartphone use among participants was more than 7 hours per day, suggesting that their blurred vision may be indirectly linked to prolonged screen time. Similarly, Liu et al. (2020) found that nearsightedness, or myopia, characterised by blurry vision, was associated with late-night sleep habits.

Regarding the burning eye sensation, a significant association was observed in the present study, consistent with findings from the previous studies by Chu et al. (2023) and An & Kim (2022), who reported that the majority of their participants experienced burning eye sensation after using the smartphone for more than 4 hours daily or having a sleep time of less than 5 hours per day. Furthermore, regarding worsening eyesight symptoms, a significant association was also observed in the present study. A similar observation was reported by Dhafira et al. (2023); however, this previous study did not find a significant association between screen time and worsening eyesight (p-value=0.094; multivariate logistic regression), even though the screen time duration (6 hours) and the sample size were almost similar to those in the present study.

The results of the present study show a significant association between prolonged screen time and back pain. A previous meta-analysis by Yue et al. (2023) comprehensively examined the relationship among these variables and found that for every 1 hour of screen time, such as daily computer use, the risk of back pain increased by up to 8.2% (Yue et al., 2023). It shows that back pain may be related to poor posture when a person is using devices. A study conducted by Thamrin et al. (2023) found that prolonged use of screen devices, averaging 7 hours or more per day, is associated with a 2.19 times higher risk of

developing low back pain. An ophthalmology expert stated that lowering one's head to view a smartphone can add weight and pressure to the neck, which is considered bad posture and should be avoided by screen users (MalaysiaNow, 2021). It is also recommended to avoid lying down while watching the screen, as this can cause muscle spasms and neck, back, and shoulder pain.

CONCLUSION

This study concludes that smartphones were the most commonly used screen devices among undergraduate students at a public university in Kuantan, Pahang. The mean addiction score to screen devices, measured using the Manolis/Roberts Cell-Phone Addiction Scale (MRCPAS) on a 5-point Likert scale, was 12.81. Overall sleep quality, assessed using the global PSQI score, showed that the majority of students were good sleepers, with only 2.6% classified as having poor sleep quality. However, excessive screen time was significantly associated with several acute health effects, including sleep disturbance, dry eyes, blurred vision, a burning-eye sensation, double vision, difficulty focusing on near vision, a feeling that eyesight is worsening, and back pain.

These findings have practical implications for students' health by highlighting the importance of promoting healthy digital habits. To minimise the risk of acute health effects from prolonged screen time, several methods can be implemented, including taking regular breaks during screen time, practising the 20-20-20 eye test rules to reduce eye strain, and limiting screen time in bed. The institutions may also run an awareness campaign to offer digital well-being education and ergonomic workshops to reduce musculoskeletal symptoms.

This study has several limitations. The limited sample size restricts the generalisations of the findings and may not represent the broader population of university students in Malaysia. Although the present study achieved a 43% response rate, which is considered acceptable for online surveys, the potential for non-response bias remains and may affect the study's statistical power. Additionally, the use of a self-reported questionnaire may have introduced response bias or inaccurate reporting. Future studies should aim to establish more precise definitions and standard guidelines for "excessive screen time" among adults. They should also consider incorporating additional behavioural, environmental and psychological factors to provide a more comprehensive understanding of the effects of screen use.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Dar al-Hikmah Library, International Islamic University Malaysia (IIUM) Kuantan, for providing access to essential databases that supported the completion of this study, as well as to all participants for their valuable contributions. This research did not receive any external funding. The authors declare no conflict of interest and contributed equally to this work.

REFERENCES

Agarwal, R., Tripathi, A., Khan, I. A., & Agarwal, M. (2022). Effect of increased screen time on eyes during COVID-19 pandemic. *Journal of family medicine and primary care, 11*(7), 3642–3647. https://doi.org/10.4103/jfmpc.jfmpc_2219_21

Albalawi, E. D., Alswayed, S. K., Aldharman, S. S., Alshangiti, A. Y., Alhussein, G. A., & Alamawi, H. O. (2023). The Association of Screen Time, Sleep Quality, and Dry Eye Among College Students in Saudi Arabia. *Cureus*, *15*(4), e37533. https://doi.org/10.7759/cureus.37533

AlQarni, A. M., AlAbdulKader, A. M., Alghamdi, A. N., Altayeb, J., Jabaan, R., Assaf, L., & Alanazi, R. A. (2023). Prevalence of Digital Eye Strain Among University Students and Its Association with Virtual Learning During the COVID-19 Pandemic. *Clinical ophthalmology (Auckland, N.Z.), 17*, 1755–1768. https://doi.org/10.2147/OPTH.S406032

Alshoaibi, Y., Bafil, W., & Rahim, M. (2023). The effect of screen use on sleep quality among adolescents in Riyadh, Saudi Arabia. *Journal of family medicine and primary care, 12*(7), 1379–1388. https://doi.org/10.4103/jfmpc.jfmpc_159_23

An, Y., & Kim, H. (2022). Sleep disorders, mental health, and dry eye disease in South Korea. *Sci Rep* 12, 11046. https://doi.org/10.1038/s41598-022-14167-0

Ataş, A., & H.Çelik, B. (2019). Smartphone Use of University Students: Patterns, Purposes, and Situations. *Malaysian Online Journal of Educational Technology*. 7. http://dx.doi.org/10.17220/mojet.2019.02.004

Bharadwaj, S. (2021). *Screen time and its impact on health*. Retrieved from

https://www.researchgate.net/publication/351048905_Screen_time_and_its_impact_on_health

Biaani, B. L. G., Palència, L., Puig-Ribera, A., Bartoll, X., & Pérez, K. (2020). Does adult recreational screen-time sedentary behavior have an effect on self-perceived health?. *Public health in practice (Oxford, England)*, *1*, 100055. https://doi.org/10.1016/j.puhip.2020.100055

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J. P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., Lambert, E., … Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British journal of sports medicine, 54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

Buysse, D. J., Reynolds, C. F., 3rd, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry research, 28(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Cain, N., & Gradisar, M. (2010). Electronic media use and sleep in school-aged children and adolescents: A review. *Sleep medicine*, *11*(8), 735–742. https://doi.org/10.1016/j.sleep.2010.02.006

Ccami-Bernal, F., Soriano-Moreno, D. R., Romero-Robles, M. A., Barriga-Chambi, F., Tuco, K. G., Castro-Diaz, S. D., Nuñez-Lupaca, J. N., Pacheco-Mendoza, J., Galvez-Olortegui, T., & Benites-Zapata, V. A. (2023). Prevalence of computer vision syndrome: A systematic review and meta-analysis. *Journal of optometry, 17*(1), 100482. https://doi.org/10.1016/j.optom.2023.100482

Chu, G. C. H., Chan, L. Y. L., Do, C. W., Tse, A. C. Y., Cheung, T., Szeto, G. P. Y., So, B. C. L., Lee, R. L. T., & Lee, P. H. (2023). Association between time spent on smartphones and digital eye strain: A 1-year prospective observational study among Hong Kong children and adolescents. *Environmental science and pollution research international*, *30*(20), 58428–58435. https://doi.org/10.1007/s11356-023-26258-0

Dhafira, F., Prihatningtias, R., Trilaksana, N. & Maharani, M. (2023). Sleep Quality and Screen Time as The Most Influential Factor of Computer Vision Syndrome. *Jurnal Promosi Kesehatan Indonesia*. 18. 73-78. https://doi.org/10.14710/jpki.18.2.73-78

Dwajani, S., Lavanya, R., Abhishek, R. S., & Praveena, A. S. (2021). Effect Of Increased Screen Time In Undergraduate Students During Covid-19 Pandemic-A Survey-Based Study. *Int J Recent Sci Res*. 11(12), pp.40252-40258. http://dx.doi.org/10.24327/ijrsr.2020.1112.5659

Hale, L., Kirschen, G. W., LeBourgeois, M. K., Gradisar, M., Garrison, M. M., Montgomery-Downs, H., Kirschen, H., McHale, S. M., Chang, A. M., & Buxton, O. M. (2018). Youth Screen Media Habits and Sleep: Sleep-Friendly Screen Behavior Recommendations for Clinicians, Educators, and Parents. *Child and adolescent psychiatric clinics of North America, 27*(2), 229–245. https://doi.org/10.1016/j.chc.2017.11.014

Hanapi, N. N. M., Ong, S. C., Ooi, G. S., & Daud, N. A. A. (2021). Sleep quality and quality of life among Malaysian pharmacy undergraduate students. *Malaysian Journal of Pharmaceutical Sciences, 19*(1), 65–75. https://doi.org/10.21315/mjps2021.19.1.5

Insana, S. P., Hall, M., Buysse, D. J., & Germain, A. (2013). Validation of the Pittsburgh Sleep Quality Index Addendum for posttraumatic stress disorder (PSQI-A) in U.S. male military veterans. *Journal of traumatic stress*, *26*(2), 192–200. https://doi.org/10.1002/jts.21793

Kaur, K., Gurnani, B., Nayak, S., Deori, N., Kaur, S., Jethani, J., Singh, D., Agarkar, S., Hussaindeen, J. R., Sukhija, J., & Mishra, D. (2022). Digital Eye Strain- A Comprehensive Review. *Ophthalmology and therapy, 11*(5), 1655–1680. https://doi.org/10.1007/s40123-022-00540-9

Kaye, L., Orben, A., A Ellis, D., C Hunter, S., & Houghton, S. (2020). The Conceptual and Methodological Mayhem of "Screen Time". *International journal of environmental research and public health, 17*(10), 3661. https://doi.org/10.3390/ijerph17103661

Khan, F. M., Shahid, M. H., Nasir, M., Karamat, A., & Ch Abdullah. (2021). Screen Time Related Musculoskeletal Symptoms among Adolescents: A Cross-Sectional Survey in Twin Cities. *Pakistan Journal of Public Health, 11*(3), 132–135. https://doi.org/10.32413/pjph.v11i3.527

Khan, W. A., N. Shah, M., Al-Ghamdi, M. A., Mohammed Al-Zahrani, A. K., Al-Khathami, A. M., Al-Hutairishi, A. M., Badri, H. M., Natto, H. A., Elamin, M. O., Milibari, A. A. & Ekram, R. (2025). A Cross-Sectional Survey

Assessing Mental Health, Fatigue and Sleep Among Male Medical Students in Western Saudi Arabia During and After Examination Periods. *Journal of Pioneering Medical Sciences, 14*(2), 59-64. https://doi.org/10.47310/jpms2025140210

Lema, A. K., & Anbesu, E. W. (2022). Computer vision syndrome and its determinants: A systematic review and meta-analysis. *SAGE Open Medicine, 10*, 20503121221142402. https://doi.org/10.1177/20503121221142402

Liu, X. N., Naduvilath, T. J., Wang, J., Xiong, S., He, X., Xu, X., & Sankaridurg, P. R. (2020). Sleeping late is a risk factor for myopia development amongst school-aged children in China. *Scientific reports*, *10*(1), 17194. https://doi.org/10.1038/s41598-020-74348-7

Lucchetti, L. (2023). *Migraine and headaches: Screen time and more*. MedicalNewsToday. https://www.medicalnewstoday.com/articles/screen-headaches-and-migraines#why-do-screens-cause-them

Malaysia ranks 3rd in world for spending most time online. (2023, November 2). *New Straits Times*. Retrieved from https://www.nst.com.my/lifestyle/bots/2023/11/974088/malaysia-ranks-3rd-world-spending-most-time-online

Melihat skrin terlalu lama beri kesan buruk pada mata dan postur. (2021, January 27). *MalaysiaNow*. Retrieved from https://www.malaysianow.com/my/news/2021/01/27/melihat-skrin-terlalu-lama-beri-kesan-buruk-pada-mata-dan-postur

Meghji, K., Talpur, M., Khan, A., Fatima, H., Memon, U., & Kewlani, M. (2025). Blue screen blue mood: Investigating the association of daily screen time with sleep quality and mental distress in students. *Journal of Rehman Medical Institute*. Advance online publication.

Mohd Kamaruzihan, N. Q. & Soe, M. K. (2023). A comparative study: impact of screen time on sleep quality among university students and school children. *Journal of Pharmacy, 3*(1). 75-85. https://10.31436/jop.v3i1.176

Montagni, I., Guichard, E., Carpenet, C., Tzourio, C., & Kurth, T. (2016). Screen time exposure and reporting of headaches in young adults: A cross-sectional study. *Cephalalgia: an international journal of headache, 36*(11), 1020–1027. https://doi.org/10.1177/0333102415620286

Park B. K. (2020). The Pittsburgh Sleep Quality Index (PSQI) and Associated Factors in Middle-school Students: A Cross-sectional Study. *Child health nursing research, 26*(1), 55–63. https://doi.org/10.4094/chnr.2020.26.1.55

Perrault, A. A., Bayer, L., Peuvrier, M., Afyouni, A., Ghisletta, P., Brockmann, C., Spiridon, M., Hulo Vesely, S., Haller, D. M., Pichon, S., Perrig, S., Schwartz, S., & Sterpenich, V. (2019). Reducing the use of screen electronic devices in the evening is associated with improved sleep and daytime vigilance in adolescents. *Sleep, 42*(9). https://doi.org/10.1093/sleep/zsz125

Poza, J. J., Pujol, M., Ortega-Albás, J. J., Romero, O. (2022). Melatonin in sleep disorders (Melatonina en los trastornos de sueño). *Neurologia*, *37*(7), 575–585. https://doi.org/10.1016/j.nrl.2018.08.002

Roberts, J. A., Yaya, L. H., & Manolis, C. (2014). The invisible addiction: cell-phone activities and addiction among male and female college students. *Journal of behavioral addictions*, *3*(4), 254–265. https://doi.org/10.1556/JBA.3.2014.015

Ross, R., Chaput, J. P., Giangregorio, L. M., Janssen, I., Saunders, T. J., Kho, M. E., Poitras, V. J., Tomasone, J. R., El-Kotob, R., McLaughlin, E. C., Duggan, M., Carrier, J., Carson, V., Chastin, S. F., Latimer-Cheung, A. E., Chulak-Bozzer, T., Faulkner, G., Flood, S. M., Gazendam, M. K., Healy, G. N., … Tremblay, M. S. (2020). Canadian 24-Hour Movement Guidelines for Adults aged 18-64 years and Adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. *Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme, 45*(10 (Suppl. 2)), S57–S102. https://doi.org/10.1139/apnm-2020-0467

Seguí, M.delM., Cabrero-García, J., Crespo, A., Verdú, J., & Ronda, E. (2015). A reliable and valid questionnaire was developed to measure computer vision syndrome at the workplace. Journal of clinical epidemiology, 68(6), 662–673. https://doi.org/10.1016/j.jclinepi.2015.01.015

Shah, N., & Hameed, S. (2023). Muscle Contraction Tension Headache. In *StatPearls*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK562274/

Thamrin, V. D., Tanjung, J. R., Haryono, I. R., & Prastowo, N. A. (2023). The Association between Screen Time and Low Back Pain among Male College Students. *Sport Mont,* *21*(2), 47-51. https://doi.org/10.26773/smj.230707

World Health Organization. (2019). Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age. World Health Organization. https://iris.who.int/handle/10665/311664.

Wu, M. J., Zhao, K., & Fils-Aime, F. (2022). Response rates of online surveys in published research: A meta-analysis. *Computers in Human Behavior Reports*, *7*. https://doi.org/10.1016/j.chbr.2022.100206

Yue, C., Wenyao, G., Yujia, L., Ya, X., Zhuying, S., Yizheng, Z., Linlin, Z., Jinxin, C., & Xingqi, W. (2023). Dose-response relationship between daily screen time and the risk of low back pain among children and adolescents: A meta-analysis of 57,831 participants. *Environmental Health and Preventive Medicine, 28*, Article 64. https://doi.org/10.1265/ehpm.23-00177

Yıldız, Ş., Bazancir Apaydın, Z., & Erbahçeci, F. (2023). The Analyses of the Relationship between Physical Activity, Musculoskeletal System Problems, Sleep, and Screen Exposure Time in University students during the Distance Learning Process. Journal of Basic and Clinical Health Sciences, 7(1), 64-74. https://doi.org/10.30621/jbachs.1059276

ORIGINAL ARTICLE

Intraocular Pressure Changes During Valsalva Manoeuvre in Young and Middle-Aged Adult Populations

Firdaus Yusof[1,2], Huzaifah Abdul Hamid[1], Noor Azian Hamdan[1], Raja Fatihah Raja Bidin[1], Nur Raihan Esa[1], Aida Hazira Ruslan[1], Syazwan Hafiz Mohd[1,*]

[1]Department of Optometry and Visual Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Bandar Indera Mahkota, 25200 Pahang, Malaysia.
[2]Integrated Omics Research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Pahang, Malaysia.

Keywords:
Intraocular pressure (IOP); Valsalva manoeuvre (VM); dynamic IOP; IOP fluctuation

ABSTRACT

Background: Intraocular pressure (IOP) is a critical physiological parameter that maintains the structural integrity and function of the eye. Understanding its dynamic behaviour under transient physiological stressors is essential for comprehensive ocular assessment. The Valsalva manoeuvre (VM) acutely increases central and episcleral venous pressures, providing a natural model for studying short-term IOP fluctuations. This study aimed to investigate and compare dynamic IOP changes during the VM between young and middle-aged adults to establish age-specific response patterns. **Methods:** A cross-sectional study was conducted among 42 healthy participants, equally divided into a young adult group (18–25 years) and a middle-aged adult group (40–65 years). IOP was recorded at four time points: pre-VM (0 s), intra-VM (10 s and 20 s) and post-VM (900 s). The VM was standardized by maintaining an expiratory pressure of 40 mmHg for 20 seconds using an aneroid manometer. Repeated-measures ANOVA with Bonferroni-corrected paired t-tests assessed IOP changes within each group, while independent-samples t-tests compared responses between age groups. **Results:** Both age groups exhibited a significant transient IOP elevation, peaking at 10 s during the VM before declining toward baseline. In young adults, mean IOP increased from 14.48 ± 3.34 mmHg (pre-VM) to 18.05 ± 3.53 mmHg (10 s), while in middle-aged adults, it rose from 14.00 ± 3.05 mmHg to 17.81 ± 3.06 mmHg (10 s). By 900 s post-VM, IOP had returned to near-baseline levels in both groups. No statistically significant inter-group difference in IOP change was found across any VM phase ($p > 0.05$). **Conclusion:** The transient IOP elevation observed in both groups reflects the physiological mechanism in which the VM raises episcleral venous pressure, temporarily impeding aqueous humour outflow. The rapid recovery of IOP to baseline indicates effective ocular and systemic autoregulation following acute venous congestion. The absence of a significant age-related difference suggests that hemodynamic defence and ocular pressure regulatory mechanisms remain comparably efficient between young and middle-aged adults.

INTRODUCTION

Intraocular pressure (IOP) is a fundamental physiological parameter essential for maintaining the structural integrity and optical function of the eye (Fautsch & Johnson, 2006). Its homeostasis depends on the balance between aqueous humour production and outflow resistance. Even minor deviations from this equilibrium can have significant clinical implications, as elevated IOP remains the most important modifiable risk factor for primary open-angle glaucoma, the leading cause of irreversible blindness globally (Asrani et al., 2024). While routine clinical evaluation emphasizes static IOP measurement, emerging evidence highlights the importance of understanding IOP's dynamic behaviour under various physiological and environmental stressors, as transient fluctuations may contribute to retinal ganglion cell damage and glaucoma progression (Liu et al., 2025; Rabiolo et al., 2024).

Among these stressors, the Valsalva manoeuvre (VM) is particularly relevant as a physiological model for acute IOP elevation. The VM involves a forced expiratory effort against a closed glottis, resulting in a transient rise in intrathoracic and intra-abdominal pressures, which in turn elevate central and episcleral venous pressures (Montes et al., 2016; Kimura et al., 2011; Aleksandrova et al., 2007). Because aqueous humour drainage requires IOP to exceed episcleral venous pressure (EVP), any acute rise in EVP can transiently impede outflow, leading to an elevation in IOP (Kazemi et al., 2025). This mechanism has been observed in both physiological contexts, such as resistance training

* *Corresponding author.*
E-mail address: syazhafiz@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

or Valsalva-like straining, and clinical settings involving venous congestion (Lara et al., 2023; Saha et al., 2022).
Previous studies have demonstrated the VM's reproducible effects on IOP and ocular blood flow (Aykan et al., 2010; Brody et al., 1999). However, comparative data across age groups remain scarce. Physiological ageing may alter autonomic regulation, vascular compliance and ocular biomechanical properties potentially influencing the magnitude and recovery of VM-induced IOP changes (Cui et al., 2024). Understanding whether these responses differ between younger and middle-aged adults is crucial, as it may reveal early age-related differences in ocular hemodynamic resilience.

Therefore, this study aimed to (1) characterize the dynamic changes in IOP during the Valsalva manoeuvre in young and middle-aged adults, and (2) compare the response profiles between these two age groups. The findings are expected to contribute to the growing body of evidence on dynamic IOP assessment and enhance clinical understanding of how acute venous stress influences ocular pressure regulation across different stages of adulthood.

MATERIALS AND METHODS

Study Design

This cross-sectional study aimed to examine IOP fluctuations during the VM among healthy young and middle-aged adults. IOP measurements were recorded before (pre-VM), during (intra-VM) and after (post-VM) the VM. Mean values at each phase were compared within and between groups.

Participants

A total of 42 healthy volunteers were recruited using purposive sampling. Participants were equally divided into two age-based groups: young adults (18–25 years) and middle-aged adults (40–65 years). Both male and female participants were included. Demographic data, refractive error status and central corneal thickness (CCT) were collected to verify homogeneity across groups and to exclude potential confounding factors (Table 1).

All participants were non-smokers, free from systemic diseases and had no contraindications to performing the VM. None had a history of ocular trauma, ocular surgery, glaucoma or any ocular pathology. Contact lens wearers and individuals on ocular medications were excluded. Only eyes with normal CCT (540 ± 30 µm) and baseline IOP values within 10–21 mmHg were included for analysis.

Ethical approval for the study was obtained from the IIUM Research Ethics Committee (IREC ID No: 85 IIUM/ 310/G/13/4/4-199- KAHS 22) and all procedures adhered to the tenets of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to data collection.

Table 1: Demographic and baseline ocular characteristics of young and middle-aged adult participants.

	Young adults N=21 (mean±SD)	Middle-aged adults N=21 (mean±SD)	*p*-value*
Mean age (years)	22.33 ± 1.24	49.05 ± 6.24	<0.001
Refractive error (DS)	-0.42 ± 0.73	0.12 ± 1.11	0.063
CCT (µm)	555.67 ± 36.91	542.00 ± 22.65	0.155

*Values are presented as mean ± standard deviation (SD). Between-group comparisons were analyzed using the Mann–Whitney U test.

Instrumentation and Materials

IOP was measured using the AccuPen handheld tonometer (Accutome, Pennsylvania, USA), which offers portability and reliability in clinical research (Horowitz et al., 2004). Prior to measurement, proparacaine hydrochloride 0.5 % (Alcon, Texas, USA) was instilled to achieve topical anaesthesia.

To standardize the Valsalva manoeuvre, participants exhaled through a sterile mouthpiece connected to an aneroid manometer (Ruicare, Shanghai, China), allowing real-time monitoring of expiratory pressure. The Oculus Pentacam (Oculus, Wetzlar, Germany) was used to measure CCT in a dark room to minimize fixation variability and light-induced accommodation.

Procedures

Pre-VM Ocular Assessment

CCT was measured with the participant's head stabilized on the chin and forehead rest. Each scan captured 100 Scheimpflug images over 2 s using the automatic release mode to minimize operator bias. Participants maintained fixation on a central red target to ensure alignment consistency.

Baseline IOP Measurement

The tonometer probe was fitted with a sterile latex tip and calibrated to zero prior to each measurement. The probe was gently tapped perpendicularly on the corneal apex and the mean of three valid readings was recorded as the

baseline IOP.

Valsalva Manoeuvre and Intra-VM IOP Measurement

Participants were seated upright and instructed to take a deep inspiration, form a tight seal around the mouthpiece with their lips and exhale forcefully to maintain an expiratory pressure of 40 mmHg for 20 s. Verbal feedback ensured steady pressure. IOP was measured at 10 s and 20 s while the target pressure was sustained, following methods adapted from prior physiological IOP studies (Aykan et al., 2010).

Post-VM IOP Measurement

Following the VM, participants rested for 900 s (15 min) in a seated position. A final IOP measurement was taken using the same procedure. This post-VM measurement represented IOP recovery following venous pressure normalization.

Data Analysis

Data were analyzed using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA). The normality of the data distribution was confirmed using the Shapiro–Wilk test. Descriptive statistics were reported as mean ± standard deviation (SD).

Within-group comparisons across the four VM phases (pre-VM = 0 s, intra-VM = 10 s, intra-VM = 20 s, post-VM = 900 s) were analyzed using repeated-measures analysis of variance (RM-ANOVA). Bonferroni-corrected paired t-tests were applied as post-hoc analyses to determine specific phase differences. Between-group comparisons (young vs. middle-aged) were analyzed using independent-samples t-tests at each time point. The significance level was set at $\alpha = 0.05$.

RESULTS

IOP Change During the VM in Young and Middle-Aged Adults

Young Adult Group

The mean baseline (pre-VM) intraocular pressure (IOP) among young adults was 14.48 ± 3.34 mmHg. A significant transient increase was observed during the VM, peaking at 18.05 ± 3.53 mmHg at the 10-second mark. This represented an approximate 24.6 % rise from baseline.

IOP decreased to 14.05 ± 3.07 mmHg at 20 s and further stabilized to 14.29 ± 2.69 mmHg at 900 s post-VM, returning to near-baseline values. Repeated-measures ANOVA revealed a significant effect of time on IOP, $F(3, 60) = 9.72$, $p < 0.001$. Bonferroni-corrected post-hoc tests indicated that IOP at 10 s differed significantly from pre-VM (0 s), intra-VM (20 s) and post-VM (900 s) measurements ($p = 0.001$, $p < 0.001$ and $p < 0.001$, respectively) (Figure 1).

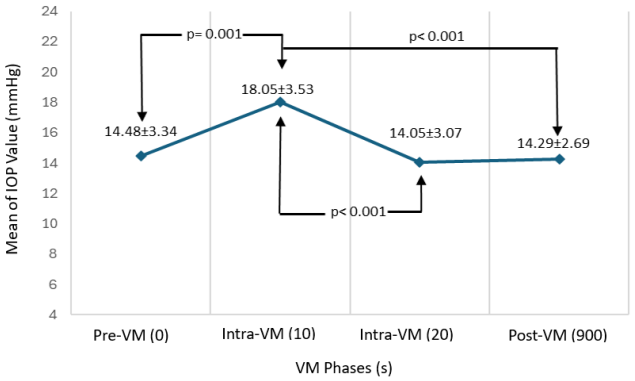


Figure 1: Mean (±SD) IOP changes across VM phases in young adults. *P*-value indicates significant differences as measured using Bonferroni-corrected paired t-tests.

Middle-Aged Adult Group

In middle-aged adults, baseline IOP was 14.00 ± 3.05 mmHg, which increased significantly to 17.81 ± 3.06 mmHg at 10 s of VM ($p = 0.001$). The mean IOP then decreased to 13.90 ± 3.19 mmHg at 20 s and 13.33 ± 2.72 mmHg at 900 s. The overall time effect was significant, $F(3, 60) = 8.83$, $p < 0.001$. Bonferroni post hoc tests indicated that only the 10-s phase differed significantly from all other phases ($p \leq 0.001$) (Figure 2).

These findings indicate a reproducible transient elevation in IOP during early VM, followed by rapid normalization in both age groups.

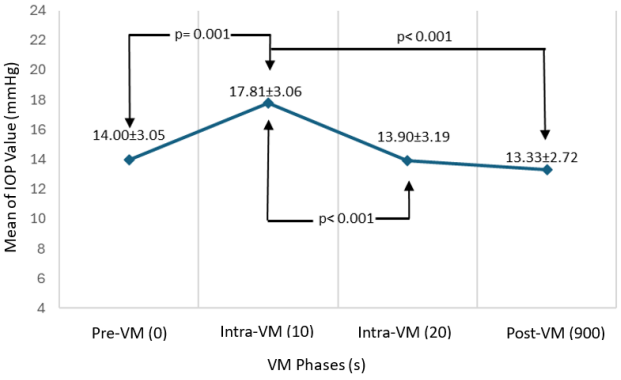


Figure 2: Mean (±SD) IOP changes across VM phases in middle-aged adults. *P*-value indicates significant differences as measured using Bonferroni-corrected paired t-tests.

Comparison of IOP Changes Between Age Groups

Figure 3 shows the mean IOP at each VM phase between the two groups. Both exhibited a similar trend: a sharp rise at 10 s, followed by a decline at 20 s and recovery at 900 s. The magnitude of IOP elevation did not differ significantly between groups at any phase (p > 0.05 for all comparisons).

At 10 s, the mean IOP was 18.05 mmHg in young adults and 17.81 mmHg in middle-aged adults, with no statistically significant difference (t(40) = 0.21, p = 0.836)(Table 2). Likewise, baseline and recovery IOP values were comparable, confirming that age did not significantly influence the amplitude or recovery of the VM-induced IOP response. RM-ANOVA indicated a significant effect of time within both groups (p < 0.001), but no significant interaction between time and age group (p > 0.05).

Table 2: Comparison of mean IOP between young and middle-aged adults across the four VM phases.

Parameter	Young adults (mean±SD, mmHg)	Middle-aged adults (mean±SD, mmHg)	*p*-value* (between groups)
Pre-VM (0 s)	14.48 ± 3.34	14.00 ± 3.05	0.701
Intra-VM (10 s)	18.05 ± 3.53	17.81 ± 3.06	0.816
Intra-VM (20 s)	14.05 ± 3.07	13.90 ± 3.19	0.883
Post-VM (900 s)	14.29 ± 2.69	13.33 ± 2.72	0.261

*Values are presented as mean ± standard deviation (mmHg). Between-group comparisons were performed using independent samples t-tests.

Figure 3 illustrates the mean IOP trend across time points for both groups, demonstrating parallel response curves.

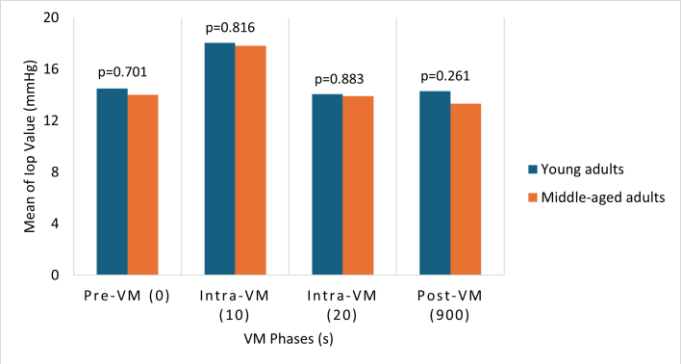


Figure 3: Comparison of IOP trajectories between young and middle-aged adults during VM phases. *P*-values from independent paired t-tests are shown for inter-group comparisons at each phase.

Both young and middle-aged adults exhibited a statistically significant but transient rise in IOP during the early phase of the VM. IOP rapidly returned to baseline within 15 minutes post-VM, demonstrating effective ocular autoregulation. No age-related differences were detected in the magnitude or recovery of IOP changes, indicating that compensatory ocular hemodynamics remain comparable between these age groups.

DISCUSSION

This study investigated the dynamic behaviour of IOP during the VM in healthy young and middle-aged adults. The main finding was a significant, transient rise in IOP during the early phase of the VM in both groups, followed by rapid recovery to baseline within 15 minutes. No statistically significant difference in the magnitude or recovery pattern of IOP was detected between age groups.

IOP Changes During the VM

The marked IOP elevation observed at 10 seconds of VM in both groups supports the well-established mechanism by which increased intrathoracic pressure leads to a rise in central and episcleral venous pressures, transiently impeding aqueous outflow and elevating IOP (Mete et al., 2016). Similar transient IOP surges during VM have been documented in earlier and recent studies, typically reaching 20–30 % above baseline before returning to normal within minutes (Lara et al., 2023; Saha et al., 2022; Aykan et al., 2010).

The post-VM return of IOP to baseline demonstrates effective ocular autoregulation and venous pressure normalization. When the expiratory strain is released, the sudden drop in intrathoracic pressure restores venous return and triggers baroreflex-mediated cardiovascular compensation (Berlin et al., 2019; Firmino et al., 2019). These systemic adjustments, together with intrinsic ocular mechanisms, rapidly re-establish aqueous outflow and IOP homeostasis. This aligns with findings from Horowitz et al. (2004), who reported that transient systemic or venous stress does not cause sustained structural deformation in healthy eyes.

Physiological and Clinical Implications

Transient IOP spikes such as those induced by VM may be clinically relevant in patients with compromised outflow pathways or reduced scleral rigidity, such as in glaucoma

or ocular hypertension (Asrani et al., 2024; Brody et al., 1999). Although healthy eyes compensate efficiently, repeated or prolonged venous stress could contribute to microvascular stress in susceptible individuals. Recognizing these short-term IOP fluctuations is important for interpreting tonometric measurements in situations involving breath-holding, coughing or heavy lifting (Saha et al., 2022).

Furthermore, studies on resistance exercise show similar patterns of transient IOP elevation during strain phases of weightlifting, emphasizing that ocular perfusion pressure and venous return are closely linked to systemic cardiovascular responses (Lara et al., 2023). These observations reinforce the need to consider body posture and breathing effort when interpreting dynamic IOP data.

Age-Related Comparison

The finding of no significant age-related effect in this study suggests that, within the examined range, ocular and systemic compensatory mechanisms remain largely preserved. Although ageing is associated with reduced vascular elasticity and altered autonomic reflexes (Cui et al., 2024), such changes may not yet be pronounced enough to affect short-term venous pressure responses in middle adulthood. Studies of older adults aged 65 and older have reported subtle delays in recovery time or smaller pressure peaks, suggesting that age-related vascular stiffness becomes increasingly influential later in life (Nilsson, 2014; Steppan et al., 2011).

Our findings indicate that the efficiency of ocular hemodynamic regulation is maintained throughout early and middle adulthood. Future studies, including older age groups and continuous IOP monitoring, could provide deeper insight into how ageing modifies ocular compliance and venous coupling.

Limitations and Future Directions

The study's limitations include its cross-sectional design and relatively small sample size, which restrict causal inference and generalization. IOP was measured at discrete time points rather than continuously; thus, potential micro-fluctuations between the 10- and 20-second intervals were not captured. In addition, systemic hemodynamic parameters such as blood pressure and heart rate were not simultaneously recorded, which could have enriched interpretation of the ocular–systemic interaction.

Future research should employ continuous tonometry or dynamic contour tonometry synchronized with cardiovascular monitoring to delineate the complete temporal IOP profile. Including older adults and patients with glaucoma or ocular hypertension would also clarify whether autoregulatory differences emerge with age or disease.

Summary

In summary, both young and middle-aged adults demonstrated a transient, statistically significant rise in IOP during the VM, followed by rapid normalization. The comparable response patterns between age groups indicate that ocular and systemic autoregulatory mechanisms maintain robust compensation during acute venous congestion. These findings enhance the understanding of dynamic IOP behaviour and provide reference data for interpreting physiological pressure changes in clinical and research contexts.

CONCLUSION

This study demonstrated that both young and middle-aged adults exhibit a transient but statistically significant elevation in IOP during the VM, peaking at the early phase before rapidly returning to baseline. The absence of significant differences between age groups suggests that ocular and systemic autoregulatory mechanisms remain efficient across early and middle adulthood in maintaining IOP stability during acute venous stress.

Clinically, these findings provide useful normative reference data for interpreting dynamic IOP fluctuations in healthy adults and underscore the need to consider transient venous influences when performing tonometric assessments. Future studies employing continuous IOP monitoring, alongside cardiovascular and choroidal measurements, could offer deeper insight into the coupling between ocular biomechanics, venous pressure, and age-related autoregulatory adaptation.

ACKNOWLEDGEMENT

The authors express their sincere appreciation to all participants and staff of the Department of Optometry and Visual Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM), for their assistance in data collection and logistical support.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare no conflicts of interest related to this study.

DECLARATION OF AI ASSISTANCE

Portions of the manuscript's language refinement, formatting and reference verification were supported by OpenAI's ChatGPT (GPT-5 model). The authors reviewed and validated all generated text, ensured factual accuracy, and take full responsibility for the content and interpretation presented in this manuscript.

REFERENCES

Aleksandrova, N. P., Donina, Z. A., & Danilova, G. A. (2007). Effects of Central Hypervolemia on Respiratory Function. *Journal of Physiology and Pharmacology, 58(5), 9-15.*

Asrani, S. G., McGlumphy, E. J, Al-Aswad, L. A., Chaya, C. J., Lin, S., Musch, D. C., Pitha, I., Robin, A. L., Wirostko, B., & Johnson, T. V. (2024). The Relationship Between Intraocular Pressure and Glaucoma. *Progress in Retinal and Eye Research, 94, 101012.* doi: 10.1016/j.preteyeres.2024.101303

Aykan, U., Erdurmus, M., Yilmaz, B., & Bilge, A. (2010). Intraocular Pressure and Ocular Pulse Amplitude Variations During the Valsalva maneuver. *Graefes Arch Clin Exp Ophthalmol, 248(8), 1183-1186.* doi: 10.1007/s00417-010-1359-0.

Berlin, D. A., Manoach, S., Oromendia, C., & Heerdt, P. M. (2019). Automated Expiratory Ventilation Assistance Through a Small Endotracheal Tube Can Improve Venous Return and Cardiac Output. *Intensive Care Medicine Experimental, 7(1), 6.* doi: 10.1186/s40635-018-0217-y

Brody, S., Erb, C., Veit, R., & Rau, H. (1999). Intraocular Pressure Changes: The Influence of Psychological Stress and the Valsalva Maneuver. *Biological Psychology, 51(1), 43-57.* doi: 10.1016/S0301-0511(99)00012-5

Cui, X., Buonfiglio, F., Pfeiffer, N., & Gericke, A. (2024). Aging In Ocular Blood Vessels: Molecular Insights And The Role Of Oxidative Stress. *Biomedicines, 12(4), 817*. doi: 10.3390/biomedicines12040817

Fautsch, M. P., & Johnson, D. H. (2006). Aqueous Humor Outflow: What Do We Know? Where Will It Lead Us?. *Investigative ophthalmology & Visual Science, 47(10), 4181–4187*. doi: 10.1167/iovs.06-0830

Firmino, E. M. S., Kuntze, L. B., Lagatta, D. C., Dias, D. P. M., & Resstel, L. B. M. (2019). Effect of Chronic Stress on Cardiovascular and Ventilatory Responses Activated by Both Chemoreflex and Baroreflex In Rats. *Journal of Experimental Biology, 222(20), jeb204883.* doi: 10.1242/jeb.204883

Horowitz, G., Byles, J., Lee, J., & D'Este, C. (2004). Comparison of the Tono-Pen and Goldmann Tonometer for Measuring Intraocular Pressure in Patients with Glaucoma. *Clinical and Experimental Ophthalmology, 32(6), 584–589*. doi: 10.1111/j.1442-9071.2004.00907.x

Kazemi, A., Roy Chowdhury, U., Fautsch, M. P., Htoo, T. M., Wirostko, B. M., Hodge, D. O., & Sit, A. J. (2025). Relationship Between Diurnal Variations of Episcleral Venous Pressure and Intraocular Pressure. *Investigative Ophthalmology & Visual Science, 66(11), 44.* doi: 10.1167/iovs.66.11.44

Kimura, B. J., Dalugdugan, R., Gilcrease III, G. W., Phan, J. N., Showalter, B. K., & Wolfson, T. (2011). The Effect of Breathing Manner on Inferior Vena Caval Diameter. *European Journal of Echocardiography, 12(2), 120-123.* doi: 10.1093/ejechocard/jeq157

Lara, P. M., Redondo, B., Jerez-Mayorga, D., Martínez-García, D., García-Ramos, A., & Vera, J. (2023). Influence of the Body Positions Adopted for Resistance Training on Intraocular Pressure: A Comparison Between the Supine and Seated Positions. *Graefe's Archive for Clinical and Experimental Ophthalmology, 261(7), 1971–1978.* doi: 10.1007/s00417-023-06009-0

Liu, T., Cai, Y., Hu, M., Wang, Z., Liu, X., Chen, M., & Wang, K. (2025). The Impact of Intraocular Pressure Fluctuations on the Progression of Glaucoma and Associated Factors. *Advances in Ophthalmology Practice and Research, 5(2), 142–148.* doi: 10.1016/j.aopr.2025.03.002

Mete, A., Kimyon, S., Saygili, O., Evisen, A., Pamukcu, C., Ceri, S., & Gungor, K. (2016). Dynamic Changes in Optic Disc Morphology, Choroidal Thickness, Anterior

Chamber Parameters and Intraocular Pressure During Valsalva Maneuver. *Arquivos Brasileiros de Oftalmologia, 79(4), 209–213.* doi: 10.5935/0004-2749.20160061

Montes, A. M., Baptista, J., Crasto, C., de Melo, C. A., Santos, R., & Vilas-Boas, J. P. (2016). Abdominal Muscle Activity During Breathing with and without Inspiratory and Expiratory Loads in Healthy Subjects. *Journal of Electromyography and Kinesiology, 30, 143-150.* doi: 10.1016/j.jelekin.2016.07.002

Nilsson, P. M. (2014). Hemodynamic Aging as the Consequence of Structural Changes Associated with Early Vascular Aging (EVA*). Aging and Disease, 5(2), 109.* doi: 10.14336/AD.2014.0500109

Rabiolo, A., Montesano, G., Crabb, D. P., Garway-Heath, D. F., & United Kingdom Glaucoma Treatment Study Investigators (2024). Relationship Between Intraocular Pressure Fluctuation and Visual Field Progression Rates in the United Kingdom Glaucoma Treatment Study. *Ophthalmology, 131(8), 902–913.* doi: 10.1016/j.ophtha.2024.02.008

Saha, B. C. & Kumari, R. (2022). Lifestyle Modifications: A New Dimension in Glaucoma Management. *Kerala Journal of Ophthalmology, 34(3), 199-207.* doi: 10.4103/kjo.kjo_56_22

Steppan, J., Barodka, V., Berkowitz, D. E., & Nyhan, D. (2011). Vascular Stiffness and Increased Pulse Pressure in the Aging Cardiovascular System. *Cardiology Research and Practice, 2011(1), 263585.* doi: 10.4061/2011/263585

ORIGINAL ARTICLE

Trend and Distribution of Refractive Errors Among Children Attending a Private Optometry Practice in Northern Malaysia

Nadiah Rashidi[1, 2], Muhammad Afzam Shah Abdul Rahim[1, 3], Firdaus Yusof[1, 3*]

[1]Department of Optometry and Visual Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia.
[2] Pusat Optometri dan Cermin Mata Nadiah Optometry, No 77 & 79, Jalan Raja Musa Mahadi, Taman Hillview, 31350 Ipoh, Perak, Malaysia.
[3]Integrated Omics Research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Kuantan, Pahang Darul Makmur, Malaysia.

Keywords:
Myopia; high myopia; children; private optometry; Malaysia

ABSTRACT

Background: Childhood myopia is escalating worldwide, increasing lifetime risks of ocular diseases and affecting educational and psychosocial outcomes. In Malaysia, practice-based evidence from private optometry is limited, constraining planning for screening and myopia-control services. **Objective:** To describe the trend and distribution of refractive errors among children attending a private optometry practice in northern Malaysia, and to examine associations with age and gender. **Methods:** This retrospective, cross-sectional review of electronic health records for 700 first-visit pediatric patients (<18 years) at a private optometry clinic in Ipoh, northern Malaysia, between April 2021 and June 2025. Refractive status was classified by spherical equivalent (SE) in the more ametropic eye as myopia (≤−0.50 D), high myopia (≤−5.00 D), hyperopia (≥+0.50 D), or emmetropia (−0.50 D < SE < +0.50 D). Age-stratified prevalence was summarized descriptively, and associations with age group and gender were tested using chi-square analysis (α = 0.05). **Results:** Myopia predominated, affecting 76.4% (n=535 from 700) of children, with 10.7% (n=75/535) classified as high myopia. Hyperopia followed, affecting 14.6% (n=102/700), and emmetropia accounted for 9.0% (n=93/700). The association between age group and refractive error type was significant (χ^2, p<0.001): myopia increases from 54.9% in preschoolers (5–6 years) to 73.6% (7–9), 74.4% (10–12), 86.8% (13–15), and 83.5% in upper-secondary students (16–17 years). Hyperopia declined across age groups, whereas emmetropia remained relatively stable. Myopia was slightly more frequent in males (80.1%) than in females (73.5%), while hyperopia was more frequent in females (18.3%) than in males (10.0%) (χ^2, p=0.008). **Conclusion:** In this northern-Malaysian private-practice cohort, myopia was highly prevalent and increased sharply across school years, with nearly one in ten children exhibiting high myopia. Hyperopia predominated in early childhood, whereas emmetropia remained relatively stable. These practice-based findings support early pediatric vision screening and integration of evidence-based myopia-control strategies within community and private care pathways. Future research should include cycloplegic refraction, axial-length monitoring, and evaluation of environmental and behavioral risk factors, as well as cost-effectiveness analyses of clinic-integrated interventions in Malaysian settings.

INTRODUCTION

Refractive errors are the most common visual disorders in children worldwide and a leading cause of visual impairment when left uncorrected (World Health Organization [WHO], 2019). Among them, myopia has drawn particular attention due to its rising prevalence, early onset, and potential to progress into high myopia. High myopia not only impairs vision but also increases the lifetime risk of sight-threatening ocular complications, including retinal detachment, myopic maculopathy, glaucoma, and cataract (Tideman et al., 2016).

Globally, the burden of myopia has grown at an alarming pace. It is estimated that by 2050, nearly half of the world's population will be myopic, and 10% will have high myopia (Holden et al., 2016; Foster & Jiang, 2014). The prevalence is especially pronounced in East and Southeast Asia, where urban school-leaving populations have reported rates exceeding 80% (Morgan et al., 2018). This "myopia epidemic" has led to significant public health concern, with myopia now recognized as one of the leading causes of preventable visual impairment in children (Holden et al., 2016; World Health Organization, 2019).

In Malaysia, available data suggest that the prevalence of pediatric myopia is increasing in parallel with regional trends (Goh et al., 2005; Saw et al., 2006). However, most published studies have been school-based (Hussin, 2021; Ismail & Sukumaran, 2022) or hospital-based (Knight et al., 2018), which may not fully represent refractive error patterns seen in routine community optometry practice.

[1,3*] *Corresponding author.*
E-mail address: yfirdaus@iium.edu.my

Children attending hospital clinics often have more severe visual conditions, while school-based surveys provide only single-time-point data without long-term follow-up. Furthermore, substantial regional variation exists: Goh et al. (2005) reported myopia in 9.8% of children aged 7 years and 34.4% of those aged 15 years in the Gombak District; Hashim et al. (2008) found a 7.7% prevalence of refractive error among Malay primary school children in suburban Kelantan; and more recently, an urban school-based study in Wangsa Maju, Kuala Lumpur, found that nearly half (47.8%) of students exhibited refractive errors, with myopia being the most common (Ismail & Sukumaran, 2022). These findings underscore the clear urban–rural disparity in childhood myopia and highlight the absence of real-world evidence from private optometry settings.

Private optometry practices play a critical role in primary eye care delivery in Malaysia, often serving as the first point of contact for parents seeking vision assessment for their children, and are therefore uniquely positioned to provide real-world data on refractive errors across a wide spectrum of pediatric patients; however, despite this important role, published studies from private practice settings are scarce, leaving their contribution to the national myopia landscape under-documented. Understanding refractive error prevalence in such contexts is particularly relevant in urban populations, where children may experience higher educational pressures, greater near-work demands, and reduced outdoor exposure, factors identified as significant risks for myopia development and progression (Wu et al., 2018; Rose et al., 2008). Given that most existing Malaysian evidence is derived from school- or hospital-based studies, which may not reflect the true distribution of refractive errors in children presenting voluntarily for care, the present study addresses this knowledge gap by analyzing pediatric refractive data from a private optometry practice in Ipoh, an urban center in northern Malaysia.

Another important knowledge gap concerns the distribution of high myopia among Malaysian children. High myopia is not only more difficult to manage clinically (Shah et al., 2024; Eppenberger et al., 2024) but also associated with a disproportionate risk of irreversible vision loss (Haarman et al., 2020; Chung, 2025). Early detection of high myopia and its age- and gender-related distribution is therefore essential to inform prevention strategies and identify children at greatest risk.

The present study addresses these gaps by analyzing pediatric refractive error data from an urban private optometry practice in northern Malaysia. The findings are intended to inform early screening initiatives, guide clinical myopia control programs, and contribute to a more comprehensive understanding of refractive error patterns among Malaysian children.

MATERIALS AND METHODS

Study Design and Setting

A retrospective cross-sectional study was conducted at a private optometry practice in Ipoh, an urban area in northern Malaysia, which provides comprehensive primary eye care services for children and adults. Data were extracted from the clinic's electronic health record (EHR) system covering all first-visit pediatric encounters between April 2021 and June 2025. EHR-based studies have been increasingly adopted to estimate prevalence and clinical patterns in real-world practice settings (Wang et al., 2012; Schneeweiss et al., 2017). All refractions were performed by qualified optometrists using a Nidek ARK-1 binocular autorefractor, followed by subjective refinement. Ethical approval for this retrospective review was obtained from the IIUM Research Ethics Committee (IREC 2025-391).

Accommodation Control and Refraction Protocol

Cycloplegic refraction is widely regarded as the reference standard for pediatric refractive assessment because pharmacologic cycloplegia suppresses accommodation and reduces the risk of over-minus prescriptions and "missed" latent hyperopia. However, non-cycloplegic refraction remains common in primary optometry settings when cycloplegia is impractical due to regulatory restrictions, parental acceptance, time, or workflow considerations. In such situations, non-cycloplegic measurements are known to yield slightly more myopic and less hyperopic spherical equivalents, especially in younger children with active accommodation (Guo et al., 2022; StatPearls, 2024).

In this study, refraction was performed without cycloplegia, consistent with standard Malaysian private-practice procedures. Objective refraction was obtained using a binocular autorefractor (Nidek ARK-1) equipped with an internal fogging system to relax accommodation. This was followed by subjective refinement using the "maximum-plus/least-minus for best visual acuity" endpoint, prescribing the most plus power (or least minus) that maintained optimal clarity. This fogging-and-maximum-plus approach is a widely accepted method to minimize accommodative effort and avoid over-minusing when pharmacologic cycloplegia cannot be performed (Review of Myopia Management, 2023; Guo et al., 2022). Despite these measures, the absence of cycloplegic agents may result in a modest underestimation of hyperopia and

an overestimation of mild myopia. This limitation, consistent with findings from comparative studies, is acknowledged and discussed later in this manuscript (Guo et al., 2022; StatPearls, 2024).

Participants

The study population comprised 700 consecutive pediatric patients (<18 years) who attended their first eye examination during the study period. Only first-visit records were included to avoid duplication. Patients with missing refraction data, prior ocular surgery, or known ocular pathology (e.g., keratoconus, congenital cataract) were excluded. Demographic variables (age, gender, and ethnicity) were obtained from the EHR and analyzed descriptively.

Data Collection and Eye Selection

Sphere and cylinder values were extracted from the EHR, and spherical equivalent (SE) was calculated as:

$$SE = Sphere + \frac{1}{2} Cylinder$$

The more ametropic eye (the eye with the greater absolute SE) was used to classify refractive status, following standard epidemiologic practice to avoid interocular correlation bias (Saw et al., 2006; Verkicharla et al., 2020).

Definitions of Refractive Error

Refractive error categories were defined according to international standards (Morgan et al., 2018; WHO, 2019):

- Myopia: SE <−0.50 D
  - High Myopia: SE ≤−5.00 D
- Hyperopia: SE ≥+0.50 D
- Emmetropia: −0.50 D < SE < +0.50 D

Age Grouping

Participants were classified into school-age groups to correspond with Malaysian educational stages:

- Preschool: 5–6 years
- Lower Primary: 7–9 years
- Upper Primary: 10–12 years
- Lower Secondary: 13–15 years
- Upper Secondary: 16–17 years

This grouping allows clinically meaningful interpretation of refractive error trends across developmental and educational stages (French et al., 2013; Wu et al., 2018).

Statistical Analysis

Data were analyzed using IBM SPSS Statistics for Windows, Version 23.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were computed for demographic and refractive variables. Prevalence rates of myopia, high myopia, hyperopia, and emmetropia were calculated overall and stratified by age group and gender. Associations between refractive-error type and both age group and gender were examined using chi-square (χ^2) tests. Statistical significance was set at $\alpha = 0.05$. Ninety-five percent confidence intervals (CIs) for prevalence estimates were computed using the binomial approximation formula:

$$CI = p \pm 1.96 \frac{\sqrt{p-(1-p)}}{n}$$

where p represents the observed prevalence and n the sample size.

RESULTS

A total of 700 pediatric patients under the age of 18 years were included in the analysis. The mean ± SD age was 10.77 ± 3.3 years (range: 5–17 years). Of these, 389 (55.6%) were females and 311 (44.4%) were males. Participants were distributed across educational-age categories as follows: preschool, 51 (7.3%); lower primary, 242 (34.6%); upper primary, 176 (25.1%); lower secondary, 152 (21.7%); and upper secondary, 79 (11.3%).

Myopia was the predominant refractive error, affecting 535 (76.4%) children. Of these, 75 (14.0%) met the criteria for high myopia (SE ≤−5.00 D), corresponding to 10.7% of the total sample. Hyperopia and emmetropia were observed in 102 (14.6%) and 63 (9.0%) children, respectively (Table 1, Figure 2).

Table 1. Distribution of refractive errors in the study sample.

Refractive Error	Proportion (%)	n
Myopia	76.4%	535
***High Myopia**	*10.7%*	*75*
Hyperopia	14.6%	102
Emmetropia	9.0%	63

*High myopia (SE ≤−5.00 D) is a subset of the myopic group; it represents 14.0% of all myopic children and 10.7% of the total sample.

Age-related variations are summarized in Table 2 and illustrated in Figure 1. The trend showed a marked increase in myopia proportion with age, from 54.9 % in preschoolers (5–6 years) to 83.5 % in upper-secondary students (16–17 years). Among the myopic children, 75 (14.0 %) had high myopia (SE ≤−5.00 D). Hyperopia declined sharply with age, while emmetropia remained relatively stable. High myopia rose progressively from 11.8% in preschoolers to 21.7% in lower-secondary

students. A chi-square analysis revealed a significant association between age group and refractive error type (χ^2, p<0.001).

Table 2. Age-related distribution of refractive errors. L represents lower, and U represents upper. The high myopia subgroup is reported within, not in addition to, the total myopic category.

Age Group	Myopia % (n)	*High Myopia % (n)*	Hyperopia % (n)	Emmetropia % (n)
Preschool	54.9% (28)	*11.8% (6)*	33.3% (17)	11.8% (6)
L/Primary	73.6% (178)	*6.6% (16)*	18.2% (44)	8.3% (20)
U/Primary	74.4% (131)	*6.8% (12)*	14.8% (26)	10.8% (19)
L/Secondary	86.8% (132)	*21.7% (33)*	7.9% (12)	5.3% (8)
U/Secondary	83.5% (66)	*10.1% (8)*	3.8% (3)	12.7% (10)

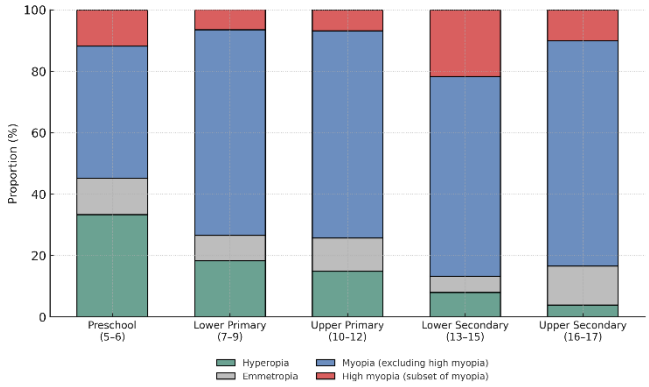


Figure 1. Age-related distribution of refractive errors

Gender differences are shown in Table 3 and Figure 2. Myopia was slightly more common in males (80.1%) than in females (73.5%). High myopia is presented as a subset of total myopia, affecting 12.9% of males and 9.0% of females. Hyperopia was more frequent among females (18.3 %) than males (10.0 %), while emmetropia proportions were similar between sexes. The overall association between gender and refractive-error type was significant (χ^2, p = 0.008).

Stratified analysis of age and gender combined revealed that the gender differences in myopia and hyperopia persisted across most age groups, with males consistently showing higher proportion of myopia and females showing a higher proportion of hyperopia. The proportion of high myopia increased with age in both genders, but no statistically significant difference was observed between males and females within individual age groups (p > 0.05).

Table 3. Gender distribution of refractive errors. The high myopia subgroup is reported within, not in addition to, the total myopic category.

Gender	Myopia % (n)	*High Myopia % (n)*	Hyperopia % (n)	Emmetropia % (n)
Male	80.1% (249)	*9.9% (31)*	10.0% (31)	10.0% (31)
Female	73.5% (286)	*9.1% (35)*	18.3% (71)	8.2% (32)

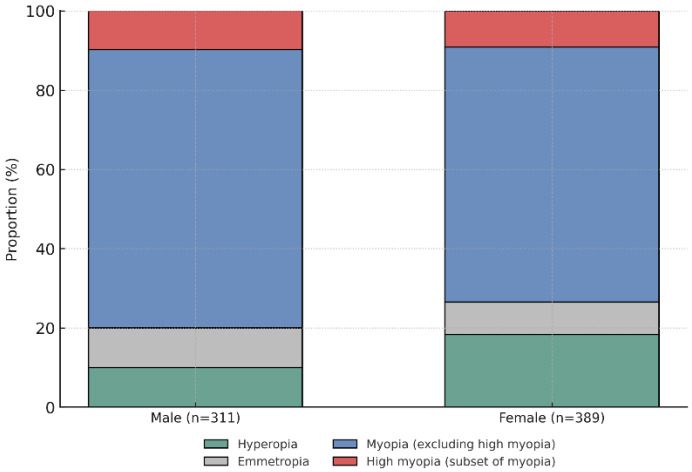


Figure 2. Gender distribution of refractive errors

DISCUSSION

This study provides practice-based epidemiological evidence on the distribution of refractive errors among Malaysian children attending a private optometry practice. Myopia was the predominant refractive error (76.4%), with high myopia present in 14.0% of myopic children (75/535) and 10.7% of the total sample. The proportion of children with myopia was higher in older age groups, ranging from 54.9% among preschoolers to 86.8% in lower-secondary students and 83.5% in upper-secondary students, while hyperopia declined correspondingly. These findings indicate a clear age-related trend and mirror patterns reported in other East and Southeast Asian cohorts (Morgan et al., 2018; Wu et al., 2018).

Comparison with previous studies

The overall prevalence in this cohort is higher than that reported in earlier Malaysian population-based studies. The National Eye Survey 1996 documented a low overall myopia prevalence (Zainal et al., 2002), but subsequent decades have shown a steady rise paralleling urbanization and educational intensification. Early regional studies, such as those by Chung et al. (1996), which found substantial rates of refractive errors among Chinese schoolchildren in Kuala Lumpur, already signaled a shift toward higher prevalence in urban areas. Goh et al. (2005) later quantified this pattern, reporting myopia in 9.8% of 7-year-olds and 34.4% of 15-year-olds in Gombak, while

Hashim et al. (2008) found a 7.7% prevalence of refractive error in suburban Kelantan. More recent urban data by Ismail and Sukumaran (2022) demonstrated that nearly half (47.8%) of schoolchildren in Kuala Lumpur exhibited refractive errors, consistent with this upward trajectory.

By contrast, studies involving rural and indigenous children consistently report far lower prevalence. Omar et al. (2019) observed only 2.4% myopia among Orang Asli children in rural Pahang, Hanisah et al. (2024) reported similar low figures among children on remote East Coast islands, and Bakar et al. (2012) found a marked difference between the more urbanized Malay and the native Iban groups. Together, these findings reveal a clear urban–rural and ethnic gradient in childhood myopia patterns in Malaysia, which aligns with differences in educational demands, near-work exposure, and outdoor activity. The markedly higher proportion observed in our urban private-practice cohort may therefore reflect both environmental and behavioural risk factors that accompany modern lifestyles.

The higher myopia trend observed here may also partly reflect self-selection bias: children attending private optometry practices typically present for symptomatic visual concerns rather than routine screening. This differs from school-based or population-level surveys (e.g., Goh et al., 2005; Ismail & Sukumaran, 2022), which include asymptomatic children. Therefore, the elevated proportion observed here likely represents the clinical burden among those actively seeking care, rather than the population prevalence.

The gender difference observed in this study, with a slightly higher proportion of myopia among males (80.1%) compared to females (73.5%), contrasts with reports from several regional studies that found a higher prevalence in females (He et al., 2004; Lin et al., 2004). This reversal may be related to sample variability or demographic differences in clinic attendance and should be examined further in population-based research.

Possible explanations and mechanisms

The age-related increase in myopia observed supports evidence that sustained near work, intensive schooling, and limited outdoor exposure accelerate the onset and progression of myopia (French et al., 2013; Wu et al., 2018; Lan et al., 2022). Environmental and behavioural influences act synergistically with biological susceptibility, resulting in excessive axial elongation and scleral remodelling (Tkatchenko et al., 2022). The proportion of high myopia (14.0% within myopia; 10.7% overall) in this sample is worrisome, as it predisposes children to retinal detachment, myopic maculopathy, and glaucoma (Tideman et al., 2016). These findings echo reports from regional cohorts (He et al., 2004; Lin et al., 2004) and are consistent with the broader epidemiologic model described by Foster and Jiang (2014), in which urbanization and educational intensity drive both the onset and progression of myopia.

Given that nearly one in ten children in this cohort had high myopia, early identification and control are crucial. Private optometry settings can play a pivotal preventive role by implementing evidence-based optical and behavioural strategies to slow progression, as supported by recent systematic reviews (Walline et al., 2020; Saw et al., 2019). Early screening before the start of primary education, parental education on outdoor activity, and myopia-control interventions (e.g., defocus spectacle lenses, orthokeratology, or low-dose atropine) are essential to mitigate progression.

Strengths and limitations

The use of electronic health records from a private optometry practice provides valuable real-world evidence rarely captured in Malaysian epidemiological literature. It complements school- and hospital-based data by reflecting children who seek primary eye care. The inclusion of a large, age-diverse sample enhances the reliability of observed trends.

However, several limitations should be acknowledged. Non-cycloplegic refraction may slightly underestimate hyperopia and overestimate mild myopia (Guo et al., 2022). Despite using internal fogging and maximum-plus endpoints, residual accommodation cannot be excluded. As this analysis was based on retrospective clinical data from a single private optometry practice, the findings represent local distribution trends rather than population prevalence. Important confounders, such as parental myopia, screen time, and outdoor duration, were unavailable in the retrospective records. Future studies incorporating cycloplegia and biometric measures will help clarify true refractive status.

Implications for practice and future directions

Despite these limitations, the findings have important implications for paediatric vision care in Malaysia. The high prevalence of myopia and its early onset emphasize the importance of routine childhood vision screening and the introduction of myopia-control interventions such as low-dose atropine, orthokeratology, and peripheral-defocus spectacle lenses in community practice. Collaboration between private practitioners and public-health services

may strengthen early detection and management, particularly in urban populations.

The escalating myopia prevalence in this and prior Malaysian cohorts underscores the need for comprehensive, multi-sectoral myopia-control strategies. Private optometrists, being the primary contact point for paediatric eye care, should be empowered to implement evidence-based interventions consistent with current management frameworks (Eppenberger et al., 2024; Shah et al., 2024; Saw et al., 2019). Integrating private practice with public health screening, promoting outdoor activities, and expanding parental awareness can help curb the trend. Longitudinal, nationwide studies incorporating cycloplegic refraction and axial length tracking are warranted to evaluate the long-term outcomes and cost-effectiveness of interventions in the Malaysian context.

Future research should include cycloplegic refraction and axial-length measurement, evaluate environmental and behavioural risk factors, and examine the longitudinal outcomes of myopia-control programmes delivered in private optometry settings. Cost-effectiveness and scalability analyses will also be essential to support sustainable national strategies for the prevention and control of myopia.

CONCLUSION

Myopia was highly prevalent among Malaysian children attending private optometric care, with rates increasing sharply across school years, and nearly one in ten were affected by high myopia. The pattern mirrors urbanization-related lifestyle changes observed across East and Southeast Asia and underscores the urgent need for early detection and intervention.

Private optometrists play a critical role in mitigating this growing public-health concern by integrating evidence-based myopia-control strategies into routine paediatric eye care. Strengthening early vision screening, promoting outdoor activity, and expanding collaboration between private and public sectors are essential steps toward sustainable myopia control in Malaysia.

DECLARATIONS

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

AI Usage Declaration

Portions of this manuscript were prepared with assistance from *ChatGPT (GPT-5, OpenAI)* under the supervision and full intellectual direction of the corresponding author. The AI tool was used solely to refine language, structure, and clarity of writing. All scientific content, analyses, and interpretations were conceived, verified, and approved by the authors.

REFERENCES

Bakar, N. F., Chen, A. H., Noor, A. R., & Goh, P. P. (2012). Comparison of refractive error and visual impairment between Native Iban and Malay in a formal government school vision loss prevention programme. *The Malaysian journal of medical sciences : MJMS*, *19*(2), 48–55.

Chung, K. M., Mohidin, N., Yeow, P. T., Tan, L. L., & O'Leary, D. (1996). Prevalence of visual disorders in Chinese schoolchildren. *Optometry and vision science : official publication of the American Academy of Optometry*, *73*(11), 695–700. https://doi.org/10.1097/00006324-199611000-00004

Chung Y. W. (2025). Myopia: a review of current concepts, association with nonophthalmological conditions, and treatment strategy in children and adolescents. Clinical and experimental pediatrics, 68(8), 554–565. https://doi.org/10.3345/cep.2025.00115

Eppenberger, L. S., Grzybowski, A., Schmetterer, L., Ang, M. (2024). Myopia control: Are we ready for an evidence based approach? *Ophthalmol Ther, 13,* 1453-1477. https://doi.org/10.1007/s40123-024-00951-w

French, A. N., Ashby, R. S., Morgan, I. G., & Rose, K. A. (2013). Time outdoors and the prevention of myopia. *Experimental Eye Research, 114*, 58–68. https://doi.org/10.1016/j.exer.2013.04.018

Foster, P. J., & Jiang, Y. (2014). Epidemiology of myopia. *Eye (London, England)*, *28*(2), 202–208. https://doi.org/10.1038/eye.2013.280

Goh, P. P., Abqariyah, Y., Pokharel, G. P., & Ellwein, L. B. (2005). Refractive error and visual impairment in

school-age children in Gombak District, Malaysia. *Ophthalmology*, *112*(4), 678–685. https://doi.org/10.1016/j.ophtha.2004.10.048

Guo, X., Shakarchi, A. F., Block, S. S., Friedman, D. S., Repka, M. X., & Collins, M. E. (2022). Non-cycloplegic compared with cycloplegic refraction in a Chicago school-aged population. Ophthalmology, 129(8), 902–910. https://doi.org/10.1016/j.ophtha.2022.02.027

Haarman, A. E. G., Enthoven, C. A., Tideman, J. W. L., Tedja, M. S., Verhoeven, V. J. M., Klaver, C. C. W. (2020). The complications of myopia: A review and meta-analysis. *Invest. Ophthalmol. Vis. Sci, 61(4),* 49. https://doi/org/10.1167/iovs.61.4.49

Hanisah J.W., Wahit K., Mustafa K., Majid A.S., Wan Yusof W.J., … & Ismail S. (2024). Refractive error and amblyopia among primary school children in remote islands of East Coast of Peninsular Malaysia. *Med J Malaysia*, *79*(5), 499-506.

Hashim, S. E., Tan, H. K., Wan-Hazabbah, W. H., & Ibrahim, M. (2008). Prevalence of refractive error in Malay primary school children in suburban area of Kota Bharu, Kelantan, Malaysia. *Annals of the Academy of Medicine, Singapore*, *37*(11), 940–946.

He, M., Zeng, J., Liu, Y., Xu, J., Pokharel, G. P., & Ellwein, L. B. (2004). Refractive error and visual impairment in urban children in southern china. Investigative ophthalmology & visual science, 45(3), 793–799. https://doi.org/10.1167/iovs.03-1051

Holden, B. A., Fricke, T. R., Wilson, D. A., Jong, M., Naidoo, K. S., Sankaridurg, P., Wong, T. Y., Naduvilath, T. J., & Resnikoff, S. (2016). Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology, 123*(5), 1036–1042. https://doi.org/10.1016/j.ophtha.2016.01.006

Ismail L. A., & Sukumaran S., (2022) Prevalence of refractive errors among school children in Wangsa Maju, Kuala Lumpur, Malaysia. *Medical hypothesis, discovery & innovation in optometry*, Vol. 3 No. 3 106-112 https://doi.org/10.51329/mehdioptometry158

Lan, W., Zhao, F., Lin, L., Li, Z., Zeng, J., Yang, Z., & Morgan, I. (2022). Refractive errors in Chinese schoolchildren: Prevalence and risk factors. *Investigative Ophthalmology & Visual Science, 63*(2), 12. https://doi.org/10.1167/iovs.63.2.12

Lin, L. L., Shih, Y. F., Hsiao, C. K., & Chen, C. J. (2004). Prevalence of myopia in Taiwanese schoolchildren: 1983 to 2000. Annals of the Academy of Medicine, Singapore, 33(1), 27–33.

Morgan, I. G., French, A. N., Ashby, R. S., Guo, X., Ding, X., He, M., & Rose, K. A. (2018). The epidemics of myopia: Aetiology and prevention. *Progress in Retinal and Eye Research, 62*, 134–149. https://doi.org/10.1016/j.preteyeres.2017.09.001

Omar, R., Wan Abdul, W. M. H., & Knight, V. F. (2019). Status of visual impairment among indigenous (Orang Asli) school children in Malaysia. *BMC public health*, *19*(Suppl 4), 543. https://doi.org/10.1186/s12889-019-6865-3

Review of Myopia Management. (2023). Ensuring an accurate refraction in children. https://reviewofmm.com/ensure-an-accurate-refraction-in-children

Saw, S. M., Goh, P. P., Cheng, A., Shankar, A., Tan, D. T., & Ellwein, L. B. (2006). Ethnicity-specific prevalences of refractive errors vary in Asian children in neighbouring Malaysia and Singapore. The British journal of ophthalmology, 90(10), 1230–1235. https://doi.org/10.1136/bjo.2006.093450

Saw, S. M., Matsumura, S., & Hoang, Q. V. (2019). Prevention and Management of Myopia and Myopic Pathology. Investigative ophthalmology & visual science, 60(2), 488–499. https://doi.org/10.1167/iovs.18-25221

Schneeweiss, S., et al. (2017). Real-world data in health care decision making: Pragmatic approaches to assess risks and benefits. *Clinical Pharmacology & Therapeutics, 102*(6), 917–929. https://doi.org/10.1002/cpt.857

Shah, R., Vlasak, N., Evans, B. J. W. (2024). High myopia: Reviews of myopia control strategies and myopia complications. *Ophthalmic Physiol Opt. 44; 1248-1260.* https://doi.org/10.1111/opo.13366

StatPearls. (2024). Cycloplegic and noncycloplegic refraction. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/books/NBK580522

Tideman, J. W., Snabel, M. C., Tedja, M. S., van Rijn, G. A., Wong, K. T., Kuijpers, R. W., Vingerling, J. R., Hofman,

A., Buitendijk, G. H., Keunen, J. E., Boon, C. J., Geerards, A. J., Luyten, G. P., Verhoeven, V. J., & Klaver, C. C. (2016). Association of Axial Length With Risk of Uncorrectable Visual Impairment for Europeans With Myopia. JAMA ophthalmology, 134(12), 1355–1363. https://doi.org/10.1001/jamaophthalmol.2016.4009

Tkatchenko, A. V., Tkatchenko, T. V., & Siegwart, J. T. (2022). *Mechanisms of scleral remodelling in myopia.* Experimental Eye Research, 219, 109047. https://doi.org/10.1016/j.exer.2021.109047

Verkicharla, P. K., Ohno-Matsui, K., & Saw, S. M. (2020). Current and predicted demographics of high myopia and an update of its associated pathological changes. *Ophthalmic and Physiological Optics, 40*(5), 402–445. https://doi.org/10.1111/opo.12778

Walline, J. J., Lindsley, K., & Vedula, S. S. (2020). *Interventions to slow progression of myopia in children.* Cochrane Database of Systematic Reviews, 1, CD004916. https://doi.org/10.1002/14651858.CD004916.pub4

Wang, S. V., Schneeweiss, S., Berger, M. L., Brown, J., de Vries, F., Douglas, I. J., … & Gagne, J. J. (2012). Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies. *Annals of Internal Medicine, 156*(11), 792–797. https://doi.org/10.7326/0003-4819-156-11-201206050-00010

World Health Organization. (2019). *World report on vision*. Geneva: World Health Organization. https://www.who.int/publications/i/item/9789241516570

Wu, P. C., Tsai, C. L., Wu, H. L., Yang, Y. H., & Kuo, H. K. (2018). Outdoor activity during class recess reduces myopia onset and progression in school children. *Ophthalmology, 125*(8), 1239–1246. https://doi.org/10.1016/j.ophtha.2017.12.011

Zainal, M., Ismail, S. M., Ropilah, A. R., Elias, H., Arumugam, G., Alias, D., Fathilah, J., Lim, T. O., Ding, L. M., & Goh, P. P. (2002). Prevalence of blindness and low vision in Malaysian population: results from the National Eye Survey 1996. The British journal of ophthalmology, 86(9), 951–956. https://doi.org/10.1136/bjo.86.9.951

ORIGINAL ARTICLE

The Effectiveness of Dichoptic Therapy in Treating Convergence Insufficiency: A Pilot Randomized Control Trial

Norsyahirah Mohamad Azmi[1], Nur Amani Izzati Mat Ghani[1], Noor Wafirah Shafee [1,2*]

[1]Department of Optometry & Visual Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia
[2] Integrated Omics Research Group (IORG), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
Convergence insufficiency; eyestrain; dichoptic mobile therapy; vergence facility; red-green filter

ABSTRACT

Background: Convergence insufficiency (CI) is a common non-strabismic binocular vision disorder that causes eyestrain, headaches, diplopia and difficulty concentrating during near tasks. It is especially prevalent among young adults who perform prolonged near work, such as university students. Vision therapy remains the mainstay for CI management. While traditional office-based and home-based therapies are effective, newer technology-driven methods such as dichoptic mobile therapy have recently emerged. However, evidence supporting their efficacy remains limited. **Aim:** To evaluate the effectiveness of dichoptic mobile therapy in improving binocular vision parameters in individuals with convergence insufficiency compared to a placebo control. **Methodology:** Fourteen volunteers aged 18–25 years with symptomatic CI were recruited and randomly assigned to an experimental group (using red–green anaglyph filters) or a placebo group (without filters). Both groups completed 20 minutes of therapy. Pre- and post-therapy assessments included visual acuity, near point of convergence (NPC), negative and positive fusional vergence (NFV, PFV), amplitude of accommodation (AA), accommodative facility and vergence facility. Data were analysed using SPSS version 20. **Result:** The experimental group demonstrated significant improvement in near PFV and vergence facility following therapy. No significant changes were observed in NPC, NFV, AA or accommodative facility. Between-group differences were not statistically significant. **Conclusion:** Dichoptic mobile therapy using red–green filters shows a potential for improving near PFV and vergence facility in individuals with convergence insufficiency. Further studies with larger sample sizes and extended intervention periods are recommended to validate these findings.

INTRODUCTION

Convergence insufficiency (CI) is one of the most common binocular vision disorders, making near tasks such as reading and computer work challenging. CI occurs when the eyes are unable to converge effectively for near vision, resulting in poor binocular vision coordination at close work. During reading, one or both eyes may drift outward, leading to symptoms such as headache, eyestrain (asthenopia), double vision (diplopia), blurred vision and difficulty concentrating or maintaining reading comprehension. CI is more prevalent among younger individuals, particularly university students, due to their increased exposure to near work, which involves prolonged screen time and reading. Reported prevalence rates of convergence insufficiency (CI) vary widely, from 1.75% to 33% (Cooper, 2012). A prevalence of CI among individuals aged 18–35 years was 27.5% (Vaishali et al., 2019), which is consistent with the 29.6% prevalence found in those aged 15–28 years (Ovenseri-Ogbomo & Eguegu, 2016).

With the rise of the Fourth Industrial Revolution (4IR), near visual demands among young individuals have increased significantly. This shift, which was accelerated by the COVID-19 pandemic and the widespread adoption of online learning , has led to prolonged use of digital devices and near tasks. Such sustained visual effort is a known risk factor for CI, as it strains the vergence and accommodative systems, leading to eyestrain, poor eye coordination and reduced concentration, which ultimately affect quality of life and academic performance (Scheiman et al., 2005). Vision therapy remains a primary intervention for managing symptoms of CI. While conventional methods have long been used, newer approaches such as dichoptic

*Corresponding author
E-mail address: wafirah@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

mobile therapy, offer a modern, technology-driven alternative.

In convergence insufficiency (CI), deficits in motor fusion and suppression limit binocular efficiency, contributing to poor vergence coordination and reduced ability to sustain convergence during near tasks (Menigite & Taglietti, 2017). Dichoptic therapy presents different visual stimuli to each eye, promoting simultaneous binocular viewing, reducing suppression, and strengthening sensory fusion (Xiao et al., 2022). By targeting these deficits, dichoptic training encourages binocular integration, enhances fusion performance, stimulates fusional reserves, and facilitates vergence adaptation, potentially addressing both the sensory and motor components underlying CI.

While conventional vision therapy has shown good outcomes in managing CI, limited research has explored the short-term effects of dichoptic mobile therapy specifically on binocular motor function. Dichoptic training may offer advantages in enhancing sensory fusion and vergence facility through simultaneous binocular stimulation using interactive digital platforms. However, its immediate clinical impact on core binocular vision parameters remains unclear. This pilot randomized controlled trial (RCT) aims to evaluate the short-term effectiveness of dichoptic mobile therapy in improving key binocular vision parameters in individuals with convergence insufficiency.

MATERIALS AND METHODS

Study design and population

This study adhered to CONSORT guidelines for pilot RCT. It was designed as a pilot RCT to evaluate the short-term effectiveness of dichoptic mobile therapy.

It also complied with the principles of the Declaration of Helsinki. Ethical approval was obtained from the International Islamic University Malaysia Research Ethics Committee (IREC) (IREC 2024-KAHS/DOVS6). G*Power version 3.1.9.3 (Heinrich Heine University Dusseldorf, Germany), a statistical power analysis tool, was used to calculate the required sample size for this study.

Fourteen volunteers aged 18–25 years with symptomatic convergence insufficiency (CI) were recruited through convenience sampling. The minimum sample size was calculated using the formula for comparing two means (Equation 1):

$$n=\frac{2\sigma^2(Z_{1-\alpha/2}+Z_{1-\beta})^2}{\Delta^2} \quad (1)$$

where σ represents the standard deviation, $Z_{1-\alpha/2}$ is the standard normal deviate for a two-tailed test at a 95% confidence level (1.96), $Z_{1-\beta}$ corresponds to a statistical power of 80% (0.84) and Δ is the minimum detectable difference between groups. Based on previous studies of convergence insufficiency therapy, an effect size (Δ) of 4 prism diopters and a standard deviation (σ) of 3.5 were assumed.

The calculation yielded a minimum of six participants per group. To compensate for potential dropouts, a 15% adjustment was applied, giving a total of 14 participants (7 per group). This sample size was deemed adequate for a pilot randomized controlled trial evaluating the short-term effect of dichoptic mobile therapy on binocular vision parameters in CI.

This pilot RCT was designed to observe and compare the short-term effects of dichoptic mobile therapy on binocular vision function in individuals diagnosed with CI. Participants were followed over a defined intervention period, allowing for the assessment of within-group changes and between-group differences, thereby strengthening the validity of the findings. Randomisation was performed using a computer-generated random number sequence, with simple random allocation into experimental and placebo groups.

Inclusion criteria for this study were comprised students from the International Islamic University Malaysia Kuantan aged 18 to 25 years, who were visual acuity of 6/6 at distance for each eye and generally healthy, not taking any medications or drugs, had never undergone refractive surgery and had score ≥16 on the Convergence Insufficiency Symptom Survey (CISS). Participants who had amblyopia, constant strabismus, nystagmus, a history of ocular trauma, vertical heterophoria greater than 1ΔD and learning disability were excluded from this study.

Data collection

Before participation, all 14 eligible participants were fully briefed about the study objectives and procedures, and written informed consent was obtained. Background information and eligibility were verified using a screening questionnaire based on inclusion and exclusion criteria. Eligible participants were then randomly assigned into two groups: Group 1 (experimental), who performed the GamE-blyopia application with red–green anaglyph filters, and Group 2 (placebo), who performed the same application without the filters.

Standardized procedures were applied for all binocular vision assessments at both baseline and post-intervention

(after 20 minutes of therapy). Near point of convergence (NPC) was assessed using a RAF ruler at 40 cm under normal room illumination. Positive and negative fusional vergence (PFV/NFV) were measured using prism bars at 40 cm with binocular viewing. Vergence facility was tested using 12Δ BO/3Δ BI prism flippers at 40 cm, while monocular and binocular accommodative facility (MAF and BAF) were evaluated using ±2.00 D lens flippers. Visual acuity (VA) was measured using a standard Snellen chart, and monocular amplitude of accommodation (AA) was assessed using the push-up method. All parameters were reassessed after the therapy session to enable comparison between pre- and post-intervention outcomes in participants with symptomatic convergence insufficiency. A summary of the flow of participants through the study procedures is presented in Figure 1.

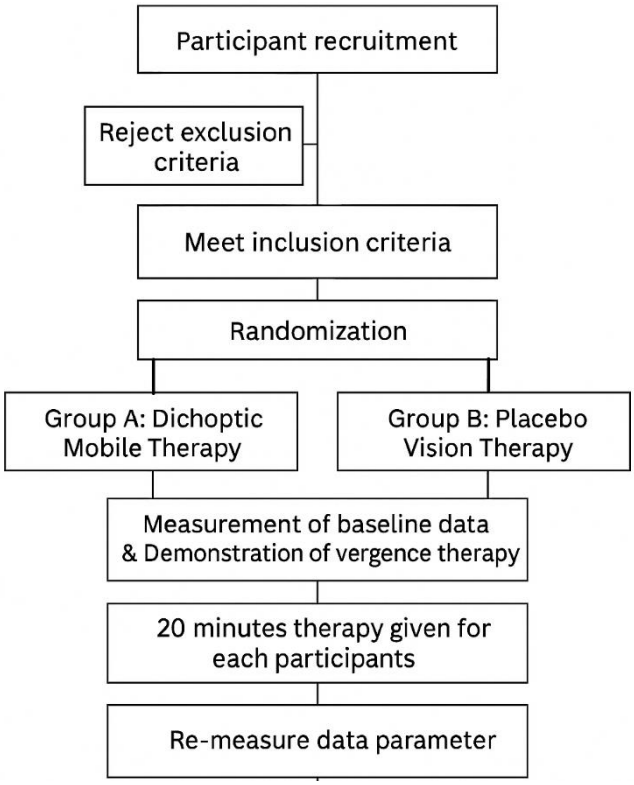


Figure 1: CONSORT flow diagram of participant recruitment, randomization, allocation, and analysis for the pilot RCT.

Assessments followed standardized order: (1) VA, (2) NPC, (3) PFV/NFV, (4) AA, (5) accommodative facility, (6) vergence facility. Testing was conducted in controlled lighting with 5-minute rest if needed. Although a 20-minute intervention does not represent the duration of conventional vision therapy, previous studies have demonstrated that short-term dichoptic and vergence-based training can produce measurable changes in binocular function and are appropriate for assessing immediate responsiveness in pilot trials (Jenewein et al., 2023).

Statistical Analysis

Data were coded and entered into Microsoft Excel, then exported to Statistical Package for the Social Sciences (SPSS) version 20 for analysis and cleaning. Means and standard deviations were calculated for the quantitative variables. Data normality was checked using the Kolmogorov-Smirnov test. Due to small sample size and non-normal distribution (Kolmogorov-Smirnov test), non-parametric tests were chosen (Wilcoxon signed-rank for within-group, Mann-Whitney U for between-group analyses). Data were screened for entry errors and outliers using boxplot and z-score (±3). No data points were excluded as no extreme outliers were detected.

RESULTS

The experimental group (with red-green filters) consisted of 8 participants (1 male, 7 females), with a mean age of 21.5 ± 1.5 years. The placebo group (without red-green filters) consisted of 6 participants (1 male, 5 females), also with a mean age of 21.5 ± 1.5 years.

Within-group comparisons demonstrated that participants in the experimental group showed significant improvements in near PFV, $p = 0.046$, and VF, $p = 0.026$, after undergoing dichoptic mobile therapy (Table 1). The increase in PFV suggests a strengthening of the convergence system and improved ability to sustain binocular fusion during near tasks. Similarly, the improvement in VF indicates enhanced flexibility and responsiveness of the vergence mechanism, reflecting better adaptation to varying vergence demands.

Other visual parameters including NPC, NFV, AA, BAF and MAF, did not show statistically significant changes following therapy. The absence of change in accommodative and convergence break/recovery measures suggests that a single 20-minute session of dichoptic training may primarily influence vergence dynamics rather than accommodative or tonic components.

In contrast, participants in the placebo group (without red-green filters) showed a significant improvement in BAF ($p = 0.039$) after therapy, while all other parameters remained statistically unchanged (Table 1). The observed increase in BAF within this group may reflect a short-term learning or practice effect from repeated testing, which can enhance accommodative response speed even in the absence of dichoptic stimulation.

Table 1: The effect of dichoptic mobile therapy on participants with symptomatic CI.

Variables	Group	N	Pre Median	Post Median	P-Value
NPC break	With R/G	8	6.00	5.50	0.157
	Without R/G	6	6.00	6.50	0.180
NPC recovery	With R/G	8	7.00	6.50	0.683
	Without R/G	6	6.50	7.00	0.180
PFV @ N	With R/G	8	17.00	20.00	0.046*
	Without R/G	6	20.50	18.00	0.500
NFV @ N	With R/G	8	18.00	18.00	0.527
	Without R/G	6	16.00	16.00	0.157
AA RE	With R/G	8	11.50	10.00	0.414
	Without R/G	6	10.00	9.50	0.564
AA LE	With R/G	8	12.00	11.50	1.000
	Without R/G	6	10.50	10.00	0.705
BAF	With R/G	8	10.00	9.50	0.257
	Without R/G	6	10.00	12.50	0.039*
MAF RE	With R/G	8	10.00	10.50	0.059
	Without R/G	6	8.50	8.00	0.461
MAF LE	With R/G	8	10.00	10.50	0.176
	Without R/G	6	7.50	9.50	0.157
VF	With R/G	8	12.50	15.00	0.026*
	Without R/G	6	9.00	10.00	0.129

Notes: R/G = Red-Green filters; NPC = Near Point of Convergence; PFV = Positive Fusional Vergence; NFV = Negative Fusional Vergence; AA = Amplitude of Accommodation; RE = Right Eye; LE = Left Eye; BAF = Binocular Accommodative Facility; MAF = Monocular Accommodative Facility; VF = Vergence Facility
*Significant level at p<0.05

Between-group comparisons using the Mann-Whitney U test revealed no statistically significant differences between the experimental and placebo groups for any of the visual parameters ($p > 0.05$ for all; Table 2). The result shows that, although the experimental group demonstrated within-group improvements in PFV and VF, these gains were not significantly greater than those observed in the placebo group.

Table 2: Comparison of visual parameters in participants with symptomatic CI after 20 minutes of dichoptic mobile therapy with and without red-green filters.

Variables	Group	N	U	P-value
NPC break	With R/G	8	16.50	0.300
	Without R/G	6		
NPC recovery	With R/G	8	17.50	0.370
	Without R/G	6		
PFV @ N	With R/G	8	11.50	0.104
	Without R/G	6		
NFV @ N	With R/G	8	20.50	0.647
	Without R/G	6		
AA RE	With R/G	8	24.00	1.000
	Without R/G	6		
AA LE	With R/G	8	19.00	0.491
	Without R/G	6		
BAF	With R/G	8	11.50	0.091
	Without R/G	6		
MAF RE	With R/G	8	23.00	0.890
	Without R/G	6		
MAF LE	With R/G	8	22.00	0.788
	Without R/G	6		
VF	With R/G	8	12.00	0.111
	Without R/G	6		

Notes: R/G = Red-Green filters; NPC = Near Point of Convergence; PFV = Positive Fusional Vergence; NFV = Negative Fusional Vergence; AA = Amplitude of Accommodation; RE = Right Eye; LE = Left Eye; BAF = Binocular Accommodative Facility; MAF = Monocular Accommodative Facility; VF = Vergence Facility; U = Mann–Whitney U Test statistic
*Significant level at p<0.05

DISCUSSION

Recent advancements in vision therapy have embraced modern technology to more effectively address binocular vision disorders. Dichoptic therapy is one such innovation: it presents different visual stimuli to each eye simultaneously (often via red-green filters), promoting binocular cooperation and reducing suppression. In the present study, dichoptic mobile therapy using red–green anaglyph filters produced significant within-group improvements in PFV and vergence facility, suggesting strengthened sensory fusion and enhanced flexibility of the motor vergence system. These findings are consistent with Li et al. (2013), who demonstrated that dichoptic training promotes binocular integration by reducing interocular suppression and enhancing cortical plasticity, even over a relatively short training duration. Similarly, Levi et al. (2015) highlighted that dichoptic stimulation can facilitate sensory fusion and improve binocular function through simultaneous stimulation of both eyes, which may explain the measurable changes in PFV observed in this pilot trial. Our findings also align with dichoptic-based interventions in amblyopia and binocular dysfunction, where improvements in sensory fusion and vergence control were reported following dichoptic training (Li et al., 2013; Levi et al., 2015).

However, unlike studies involving longer-duration therapy or multiple training sessions that demonstrated broader improvements across NPC, accommodative facility, and symptom scores (Cooper, 2012; CITT-ART Investigator Group, 2019), this study demonstrated measurable change predominantly in PFV and VF, which are known to respond more rapidly to vergence stimulation. The improvement in VF indicates increased vergence adaptability and response speed, consistent with Jenewein et al. (2023), who reported early improvements in vergence dynamics during the initial phase of therapy. The absence of significant changes in accommodative parameters and NPC in our study may be attributed to the short intervention duration, as these parameters may require cumulative training effects or repeated sessions to demonstrate meaningful change. While no between-group differences were found, the within-group improvements in PFV and VF support the potential of dichoptic mobile therapy as a short-term stimulus for enhancing sensory fusion and motor vergence function in individuals with CI.

For example, the multicenter Convergence Insufficiency Treatment Trial-Attention & Reading Trial (CITT-ART) found a mean increase in PFV of approximately 14.4Δ ($P < 0.001$) after office-based vergence/accommodative therapy in children with symptomatic CI (CITT-ART Investigator Group, 2019). Similarly, a time-course study reported that the fastest rate of improvement in NPC and PFV occurred within the first 4 weeks of therapy, with PFV improving by about 3.2Δ/week (Jenewein et al., 2023). Moreover, review evidence indicates that computer-based or dichoptic therapy in adults with CI can lead to increased PFV (e.g., mean increase of 17.7Δ in a small sample) and reduced symptoms (Cooper, 2012).

Thus, the significant improvements seen in PFV and VF in our pilot RCT suggest that the dichoptic mobile therapy may activate and enhance the vergence system's ability to adapt rapidly to fusion demands. Because VF assesses the speed and flexibility of vergence responses (cycles per minute, cpm), an increase in VF denotes improved motor fusion responsiveness, which is critical for sustained near tasks. By promoting binocular cooperation through red-green filter dichoptic stimuli, the therapy likely strengthens fusional vergence reserves and supports sensory fusion mechanisms.

To extract direct practical implications: in CI, the difficulty often lies in sustaining convergence under near demands and resisting exophoria drift. Improvements in PFV enlarge the fusional reserve capacity and enhanced VF indicates better dynamic responsiveness. Hence, our findings align well with theoretical frameworks that reinforce the role of targeted binocular stimuli in treating convergence dysfunctions.

Interestingly, the placebo group (without red-green filters) also demonstrated a significant improvement in binocular accommodative facility (BAF). This improvement may reflect a practice or learning effect, rather than a genuine change in binocular fusion or vergence mechanisms. In other words, repeated exposure to the testing procedure, increased familiarity with the tasks and greater visual engagement during the therapy session may have enhanced the speed and accuracy of the accommodative response even in the absence of dichoptic stimulation. Several studies support this phenomenon, one reliability study found significant increases in accommodative facility (both monocular and binocular) simply due to repeated testing over three consecutive weeks, suggesting a learning effect on the facility measure. Another study exploring accommodative facility training reported that facility can improve via direct training in myopes and emmetropes over short time intervals (McKenzie et al., 1987). Given these findings, the improvement in BAF in the placebo group likely represents task-familiarity rather than an effect of the dichoptic therapy, and this should be considered when interpreting the outcome (Allen et al., 2010).

Despite these within-group improvements, no statistically significant differences emerged between the experimental and placebo groups across all parameters. Importantly, this absence of between-group difference is not necessarily due to insufficient sample size (our sample met the pre-calculated minimum), but is more likely explained by the short duration of the therapy session (20 minutes). It is plausible that a longer or repeated therapy regimen would be required to produce measurable inter-group differences in binocular function.

Overall, the findings suggest that dichoptic mobile therapy using red-green filters can generate measurable short-term improvements in vergence performance (particularly PFV and VF) among individuals with CI. However, the results also indicate that sustained and broader improvements across other binocular and accommodative parameters likely require longer intervention periods or repeated sessions. Moreover, the interactive and gamified nature of the mobile platform may enhance patient engagement and adherence compared to conventional therapies, a promising feature for clinical implementation.

CONCLUSION

Recent innovations such as dichoptic mobile therapy, delivered through the GamE-blyopia application application using red-green filters, show promise in enhancing near PFV and VF among individuals with CI. This approach may serve as an effective and accessible adjunctive therapy to conventional vision training by improving binocular motor function while increasing patient motivation and compliance through its interactive design. However, as this was a short-term pilot study with a small sample size, further research with larger cohorts, extended therapy durations and long-term follow-up is needed to validate its therapeutic potential and sustained clinical efficacy.

LIMITATIONS

Although the sample size in this study was small, it met the minimum requirement based on prior sample size calculations to achieve adequate statistical power for detecting meaningful differences. Therefore, the sample size is considered sufficient for a pilot randomized controlled trial. However, as a preliminary investigation, the limited number of participants may restrict the generalizability of the findings to a broader population. In addition, the uniform duration and frequency of therapy sessions may have influenced the measured outcomes. Future studies should include larger and more diverse participant groups, extended follow-up periods and comparisons with alternative treatment modalities such as conventional orthoptic exercises or virtual reality-based vision therapy to better evaluate the long-term efficacy and clinical applicability of dichoptic mobile therapy.

ACKNOWLEDGEMENT

The author extends sincere appreciation to the Department of Optometry and Visual Sciences, International Islamic University Malaysia (IIUM), Kuantan Campus, for their support and to all participants for their involvement in this study.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author(s) used ChatGPT (OpenAI) to assist in drafting and refining the text. The author(s) subsequently reviewed, revised, and approved all content, and accept full responsibility for the final manuscript.

REFERENCES

1. Allen, P. M., Charman, W. N., & Radhakrishnan, H. (2010). Changes in dynamics of accommodation after accommodative facility training in myopes and emmetropes. *Vision Research*, *50*(10), 947–955. https://doi.org/10.1016/j.visres.2010.03.007

2. CITT-ART Investigator Group. (2019). Treatment of symptomatic convergence insufficiency in children enrolled in the convergence insufficiency treatment trial-attention & reading trial: A randomized Clinical trial. *Optometry and Vision Science*, *96*(11), 825–835. https://doi.org/10.1097/OPX.0000000000001443

3. Cooper, J. (2012). Convergence insufficiency-a major review. In *Article in Optometry-Journal of the American Optometric Association*. https://www.researchgate.net/publication/233899973

4. Jenewein, E. C., Cotter, S., Roberts, T., Kulp, M., Mitchell, G. L., Jones-Jordan, L. A., Chen, A. M., Hopkins, K., Huang, K., Amster, D., Fecho, G., Tyler, J., Meiyeppen, S., Scheiman, M., Cooper, J., Schulman, E., Hamian, K., Iacono, D., Larson, S., & Varghese, R. (2023). Vergence/accommodative therapy for symptomatic convergence insufficiency in children: Time course of improvements in convergence function. *Ophthalmic and Physiological Optics*, *43*(1), 105–115. https://doi.org/10.1111/opo.13062

5. Levi, D. M., Knill, D. C., & Bavelier, D. (2015). Stereopsis

and amblyopia: A mini-review. In *Vision Research* (Vol. 114, pp. 17–30). Elsevier Ltd. https://doi.org/10.1016/j.visres.2015.01.002

6. Li, J., Thompson, B., Deng, D., Chan, L. Y. L., Yu, M., & Hess, R. F. (2013). Dichoptic training enables the adult amblyopic brain to learn. In *Current Biology* (Vol. 23, Issue 8). Cell Press. https://doi.org/10.1016/j.cub.2013.01.059

7. McKenzie, K. M., Kerr, S. R., Rouse, M. W., & DeLand, P. N. (1987). Study of accommodative facility testing reliability. *Am J Optom Physiol Opt.*, *64*(3), 186.

8. Menigite, N. C., & Taglietti, M. (2017). Visual symptoms and convergence insufficiency in university teachers. Revista Brasileira de Oftalmologia, 76(5), 242–246. https://doi.org/10.5935/0034-7280.20170050

9. Ovenseri-Ogbomo, G. O., & Eguegu, O. P. (2016). Resultados vergenciales y disfunciones de las vergencias horizontales entre los estudiantes universitarios de primer año de Benin City, Nigeria. *Journal of Optometry*, *9*(4), 258–263. https://doi.org/10.1016/j.optom.2016.01.004

10. Scheiman, M., Lynn Mitchell, ; G, Cotter, S., & Cooper, J. (2005). A Randomized Clinical Trial of Treatments for Convergence Insufficiency in Children. Arch Opthalmol., 123, 14–24.

11. Vaishali, R., Jha, K., & Srikanth, K. (2019). Prevalence of convergence insufficiency between 18 and 35 years and its relation to body mass index. *TNOA Journal of Ophthalmic Science and Research*, *57*(1), 27. https://doi.org/10.4103/tjosr.tjosr_11_19

12. Xiao, S., Angjeli, E., Wu, H. C., Gaier, E. D., Gomez, S., Travers, D. A., Binenbaum, G., Langer, R., Hunter, D. G., & Repka, M. X. (2022). Randomized controlled trial of a dichoptic digital therapeutic for amblyopia. Ophthalmology, 129(1), 77–85. https://doi.org/10.1016/j.ophtha.2021.09.001

ORIGINAL ARTICLE

Digital Learning's Toll: A High Prevalence of Computer Vision Syndrome in University Students During Lockdown

Lim Co Shin[1], Lili Asma Ismail[1], Shah Farez Othman[2], Mohd Hafidz Ithnin[2], Mohamad Hanif Hajar Maidin[2,*]

[1]School of Optometry, Faculty of Medicine and Health Sciences, UCSI University, 56000 Kuala Lumpur, Malaysia

[2]Ophthalmic Sciences Research Group(OSReG), Department of Optometry and Visual Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
Computer Vision Syndrome (CVS), Computer Vision Symptoms Scale 17 (CVSS17), External Symptom Factors (ESF), Internal Symptom Factors (ISF), COVID-19, online learning

ABSTRACT

Background: Computer Vision Syndrome (CVS) encompasses a range of ocular and musculoskeletal symptoms, such as dry eyes, eye strain, blurred vision, neck pain, and light sensitivity, resulting from prolonged digital device use. During the COVID-19 lockdown, the shift to online learning increased screen exposure, potentially exacerbating CVS among university students. This study aimed to assess the CVS prevalence, and its associated Internal Symptoms Factors (ISF) and External Symptoms Factors (ESF) among students at UCSI University, Kuala Lumpur, during the pandemic-induced lockdown.
Method: A cross-sectional online survey was conducted using the Computer Vision Symptoms Scale 17 (CVSS17) questionnaire. **Results**: Among 150 participants (mean age 21.36 ± 1.87 years; 99 females, 51 males), the mean CVSS17 score was 34 ± 10.61, with ESF and ISF scores of 24 ± 7.07 and 15 ± 4.95, respectively. A high prevalence of CVS was observed, with 80.67% reporting moderate to severe symptoms. Specifically, 56.67% experienced moderate to severe ESF, while 83.33% reported moderate to high ISF. Female students exhibited significantly higher CVSS17, ESF, and ISF scores compared to males ($p<0.01$). **Conclusion**: The findings indicate a substantial prevalence of CVS among students during the lockdown, driven predominantly by internal symptom factors. These results highlight the need for preventive measures to mitigate CVS in digital learning environments.

INTRODUCTION

Computer Vision Syndrome (CVS), clinically termed digital eye strain, has emerged as a significant occupational and environmental health challenge in our technology driven society. This multifactorial condition arises from prolonged interaction with digital displays-including computers, tablets, and smartphones and is characterized by a constellation of visual, ocular and musculoskeletal symptoms (Chawala et al., 2019).

The pathophysiology of CVS involves three primary mechanisms; ocular surface disruption (dry eye, irritation, redness), accommodative dysfunction (eye strain, blurred vision) and ergonomics related strain (neck/shoulder pain, migrain (Wimalsundera, 2006; Parihar et al. 2021). The clinical presentation of CVS ranges from transient visual disturbances to chronic discomfort, with studies reporting 50-90% prevalence among regular device users (Bali et al., 2007), increased 2.8x risk of severe symptoms with more than 4 hours daily use (Kim et al. 2016) and significant comorbidity with cervicobrachial disorders (Berolo et al., 2011).

Unlike traditional print media, digital screens present distinct visual stressors. Letter on screens, made up of pixels usually have clear sharp central and decrease toward the edges, reading with a digital screen could be attributed to screen brightness, glare, font size, contrast, and eye to screen distance (Lema and Anbesu, 2022). Compared to the static contrast of print media, digital displays are pixel dependent. Turgut (2018) reported 40% greater accommodative demand when viewing a digital screen. Glares are more apparent with digital screen reflectance, Lee et al. (2016) reported of 2.5x increased of photophobia risk. Fixed ergonomics with digital displays cause as high as 68% musculoskeletal complaints (Parihar et al. 2021).

The COVID-19 pandemic triggered a profound global transformation, marked by widespread lockdowns and unprecedented societal disruptions. In Malaysia, multiple iterations of the Movement Control Order (MCO) were implemented over two and a half years to mitigate the crisis. These measures impacted all sectors, with education facing particularly acute challenges. Educational

*Corresponding author.
E-mail address: hanifmaidin@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

institutions worldwide were forced to rapidly transition to online learning as their primary instructional mode. While this shift enabled continuity in teaching, it also introduced new pedagogical complexities and raised concerns about health risks linked to prolonged computer use notably, computer vision syndrome (CVS).

MATERIALS AND METHODS

This cross-sectional study investigated the prevalence and symptom profile of Computer Vision Syndrome (CVS) among university students during mandatory online learning under COVID-19 lockdown conditions. The research had two primary objectives: to quantify CVS prevalence in this unique educational context, and to identify the most significant contributing symptoms and risk factors. The study employed a convenience sampling approach, with all research activities - including participant recruitment, screening, and data collection - being conducted virtually due to movement restrictions. Prior to participation, all respondents received complete study information through digital patient information sheets. Written informed consent was obtained electronically from each participant before they accessed the study questionnaire.

The required sample size was calculated a priori using G*Power version 3.1.9.3 (Heinrich Heine University Dusseldorf, Germany). Based on a two-tailed independent t-test design comparing means between two groups, we established the following parameters, statistical power (1-β) was set at 0.8, significant level (α) at a 0.05, 0.50 effect size (d) and 0.5 ratio of two groups. Analysis yielded a critical t value of 1.66, degree of freedom (df) 112, minimum total sample required 114 participants with larger group of 76 participants and 38 participants for smaller groups. This calculation followed established methodological recommendations for research (Faul et al., 2017). The effect size was selected based on comparable studies of visual symptoms in digital device users, while the allocation ratio reflected the expected gender distribution in our target population.

The study enrolled participants aged 17 years or older who reported a minimum computer usage of either four hours daily or twenty hours weekly and who provided informed consent. Individuals were excluded if they presented with active ocular inflammation or trauma, severe refractive errors exceeding ±6.00 dioptres, systemic diseases affecting ocular health, or current use of medications known to influence ocular surface integrity or visual function—including but not limited to antihistamines, beta-blockers, diuretics, psychotropic agents, and oral contraceptives. Participants with corrected refractive errors were instructed to respond to survey questions while wearing their habitual optical correction. The study protocol received ethical approval from the Institutional Review Board (ICE-2021-FMH-074) and was conducted in full compliance with the ethical principles of the Declaration of Helsinki, ensuring protection of participant rights and welfare throughout the research process.

The study employed the validated Computer Vision Syndrome Severity 17 (CVSS17) questionnaire (Gonzalez-Perez et al. 2018), a psychometrically robust instrument developed using the Rasch measurement model to objectively quantify CVS symptom severity. This 17-item tool transforms ordinal response data into interval level measurement by accounting for both item difficulty and respondent ability, thereby enabling precise evaluation of symptom severity. The questionnaire incorporates three response formats: a seven-point frequency scale (never-always) for one item, a six-point intensity scale (not at all-very much) for six items, and four-point scales (never-always or false-true) for the remaining ten items. Total scores (range: 17-53) were calculated by multiplying the sum of responses by 17 and dividing by the number of valid answers, with higher scores indicating greater severity. These scores were categorized into five severity levels (Level 1 (17-22), Level 2 (23-28), Level 3 (29-35), Level 4 (36-42), and Level 5 (43-53).

CVSS17 evaluates two subscales: the External Symptom Factor (ESF) and the Internal Symptom Factor (ISF). Each factor comprises eleven and six items, respectively. ESF assesses external factors influencing ocular symptoms, specifically dry eye symptoms associated with computer use. On the other hand, ISF examines factors related to the vision function, such as refraction and accommodation. ESF total scores are calculated based on the sum of score x 11 questions over the number of valid responses. There are five performance levels: Level 1 (score between 11 and less than 18), Level 2 (score between 18 and less than 24), Level 3 (score between 24 and less than 30), Level 4 (score between 30 and less than 35) and Level 5 (score greater than or equal to 35). ISF total scores are calculated based on the sum of score x 7 questions over the number of valid responses. Three performance levels are defined: Level 1 (score between 7 and less than 12), Level 2 (score between 12 and less than 16) and Level 3 (score greater than or equal to 16).

All statistical analyses were conducted using IBM SPSS Statistics 29.0, employing descriptive statistics to determine prevalence rates and chi-square tests (with and without Yates' correction) to examine gender-based associations, with statistical significance set at $p < 0.05$.

RESULTS

Descriptive statistics are summarised in Table 1.0. A total of 150 students participated, comprising 99 females and 51 males with mean age of 21.36 ± 1.87 years. A significant proportion of students were in their fourth year, followed by second-year, fifth year, and third year. The average CVSS17 score was 34.43 ± 6.29, indicating that 80.67% of participant experienced a moderate to severe CVS. The average ESF score was 23.68 ± 4.59 (56.67%) and the average ISF score was 14.56 ± 2.84 (83.33%).

Table 1.0: Socio-demographic Characteristics of Study Participants

Characteristic	Number, n /Score (%)
Mean age	21.36 ± 1.87 years
Gender	
i. Male	51 (34%)
ii. Female	99 (66%)
Year of study	
i. First	24 (16.00%)
ii. Second	28 (18.67%)
iii. Third	19 (12.67%)
iv. Fourth	53 (35.33%)
v. Fifth	26 (17.33%)
CVSS17 Score	34.43 ± 6.29 (80.67%)
ESF Score	23.68 ± 4.59 (56.67%)
ISF Score	14.56± 2.84 (83.33%)

Figure 1.0 shows the distribution of responses from all participants to the CVSS17 questionnaires, ranging from the most negative to the most positive items according to question number. Table 1.1 shows the distribution of CVSS17 levels among study participants.

Table 1.1: Distribution of CVSS17 Levels for Participants

CVSS17	1	2	3	4	5
Male	3	11	23	14	0
Female	2	13	22	50	12

Figure 1.1 provides a detailed breakdown of the study responses for the ESF component while Table 1.2 shows the distribution breakdown of ESF level among all participants. Figure 1.2 presents detailed responses for ISF related questions and Table 1.3 summarises the distribution breakdown of ISF levels of all participants.

Overall, 80.67% (121 of 150) participants experience CVSS17 level 3 and above. In terms of gender, 72.54% (37 of 51) male participants experience moderate to severe CVS compared to 84.84% (84 of 99) female participants. The most frequently reported symptom was under External Symptom Factor, ESF was tired eyes (ESF question A4), with 78% of participants reporting experiencing it from occasional to constant. 74.67% of participants concurred that they needed to shut their eyes to alleviate dry eye symptoms (ESF question C23), while 72.67% agreed they experienced a sensation of heaviness in their eyes at the end of working day (ESF question C16), 58.67% of participants reported heavy eye after working on the computer for a while (ESF question A17). Other symptoms reported by less than 50% of study participants included excessive blinking (ESF, question A20) at 46%, eye pain (ESF question A9) at 35.55%, a stinging sensation in the eyes (ESF question A32) at 29.33%, watery eyes (ESF question B7) at 27.33%, red eye after prolonged computer work at 22.67% and burning eye sensation (ESF question A21) at 21.33%.

Based on Table 1.2., 56.67% (85 of 150) participants experienced ESF level 3 and above. In terms of gender, 39.22% (20 of 51) male participants experience moderate to severe CVS compared to 65.65% (65 of 99) female participants.

Table 1.2: Distribution of ESF Score Level for All Participants

ESF	1	2	3	4	5
Male	9	22	20	0	0
Female	9	25	57	8	0

Under ISF questions, 68.67% agreed with the statement that they were bothered by light after extended periods of work (ISF question C24), and 68% concurred they had to exert extra effort to see clearly (ISF question C21). This is followed by eye strain (ISF, question A22) at 41.33%, followed by discomfort due to lighting (ISF question A33) at 40%, blurred text on the screen (ISF question A2) at 30%, a feeling of crossing the eyes (ISF question A28) at 25.33% and double vision (ISF question A30) at 23.33%.

Table 1.3: Distribution of ISF Score Level for All Participants

ISF	1	2	3
Male	13	24	14
Female	12	39	48

Based on Table 1.3., 83.33% (125 of 150) participants experience ISF level 2 and above. In terms of gender, 74.50% (38 of 51) male participants experience moderate to severe ISF compared to 87.87% (87 of 99) female participants.

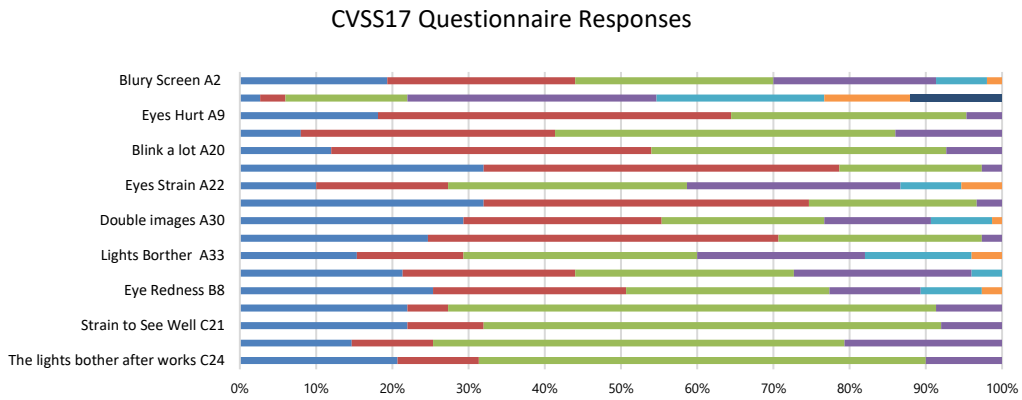


Figure 1.0: Stacked Cluster Bar Chat of CVSS17 Questionnaire Responses for All Participants

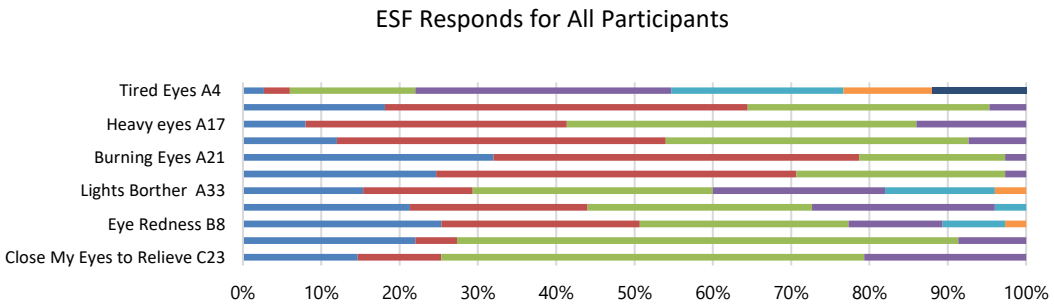


Figure 1.1: Stacked Cluster Bar Chat of CVSS17 ESF Responses for All Participants

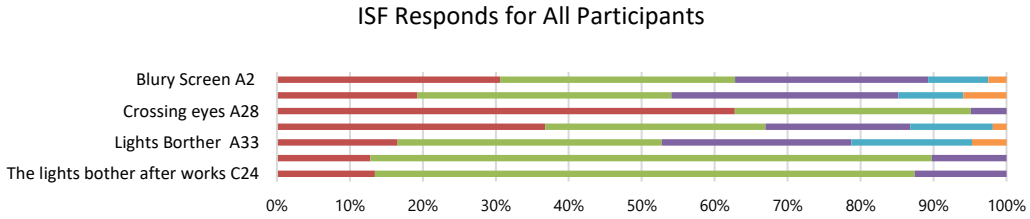


Figure 1.2: Stacked Cluster Bar Chart of ISF Responses for All Participants

Table 1.4 Comparison of CVS score for all participants and result Mann-Whitney U Test Between Genders

	CVSS17	ESF Score	ISF Score
Overall (n=150)	34.43 ± 6.29	23.68 ± 4.59	14.56 ± 2.84
Prevalence	80.67%	56.67%	83.33%
Male (n=51)	32.02 ± 5.82 Mean rank 58.87	22.04 ± 4.26 Mean rank 59.73	13.45 ± 2.68 Mean rank 58.98
Female (n=99)	35.68 ± 6.19 Mean rank 84.07	24.53 ± 4.55 Mean rank 83.63	15.13 ± 2.76 Mean rank 84.01
Z score	U=1676.50, z=3.37 $p<0.01^{\dagger\dagger}$	U=1720, z=3.20 $p<0.01^{\dagger\dagger}$	U=1682, 3.36 $p<0.01^{\dagger\dagger}$
Effect size	r=0.27	r=0.26	r=0.27

p-value, significance value; $p<0.01^{\dagger\dagger}$

Table 1.5 Significant Association Between Gender and the Symptoms Experienced by the Participants During Their Last Four Working Weeks

Symptoms	Pearson Chi-square	df	p-value
Strains to See Well C21	10.29	1	$<0.01^{\dagger\dagger}$
Blurry Letters A2	7.18	1	$<0.01^{\dagger\dagger}$
Watery Eyes B7	6.63	1	$<0.05^{\dagger}$
Tired Eyes A4	5.44^{γ}	1	$<0.05^{\dagger}$
Light bother After Works C24	5.00	1	$<0.05^{\dagger}$

$^{\gamma}$ Chi-square with Yates's correction, df, degree of freedom; p-value, significance value; $p<0.05^{\dagger}$, $p<0.01^{\dagger\dagger}$

The results presented in Table 1.4 show significant differences between genders for CVSS17 ($p<0.01$), ESF score ($p<0.01$), and ISF score ($p<0.01$). The Mann-Whitney test revealed that female participants had significantly higher CVSS17 scores (mean rank=84.07, n=99) compared to male participants (mean rank=58.87, n=51), with U=1676.50, z = 3.37, and $p < 0.01$ (two-tailed), indicating a small effect size (r = 0.27). Similarly, the ESF score for female participants (mean rank=83.63, n=99) was significantly higher than for male participants (mean rank=59.73, n=51), with U=1720, z=3.20, $p<0.01$ (two-tailed) and a small effect size (r=0.26). The ISF score showed a similar pattern, with female participants (mean rank=84.01, n=99) having significantly higher scores than male participants (mean rank=58.98, n=51), U=1682, z=3.36, $p<0.01$ (two-tailed), a small effect size (r=0.27).

The chi square test results presented in table 1.5 show that several questions revealed significant differences between male and female participants. For the ESF related questions, significant differences were observed for blurry vision on screen (question A4) with χ^2 (1, N=150) =7.18, $p<0.01$ and watery eyes, (question B7) with χ^2 (1, N=150) =6.63, $p< 0.05$). For ISF, significant differences were found for blurry letters (question A2) with χ^2 (1, N=150) =7.18, $p<0.01$, the need to strains to see well (question C21) with χ^2 (1, N=150) =10.29, $p<0.01$, and light bothering after wok (question C24) with χ^2 (1, N=150) =5.00, $p< 0.05$).

DISCUSSION

The transition to exclusive online learning during the COVID-19 lockdown provide a unique context to examine CVS prevalence among university students. Our study reveals 80.7% of respondents (n=150) experienced moderate to severe symptoms, underscoring the substantial ocular strain associated with prolong digital device use in academic settings. This finding carries important implications for both public health and educational policy as digital learning becomes increasingly prevalent.

When contextualized within the existing literature, our prevalence rate demonstrates both consistencies and divergences that merit discussion. Among Malaysian populations, our results align closely with Yuan et al. (2022) (76.7% prevalence) while being slightly lower than Reddy et al. (2013) (89.9%), likely reflecting differences in assessment methodologies (CVS-Q vs custom survey instrument) and sampling frame.

However, global literature on CVS prevalence presents intriguing variations. South Asian studies report an extreme range from 71.6% in Nepal (Sitaula and Khatri (2018) to 94.8% in Pakistan Hassan et al. (2017). Middle Eastern data shows moderately high prevalence at 72% in UAE (Shantakumari et al. 2014). Indian studies demonstrate significant methodological variation, with Logaraj et al. (2014) reporting 78.7-81.9% among professional students versus Ganne et al. (2021) documenting just 50.6% during COVID-19. These disparities likely stem from three key factors, population characteristics (medical versus general students), diagnostic criteria (validated tools vs symptom checklists) and environmental contexts (pre-pandemic vs lockdown condition).

Notably, our lockdown-era findings remain consistent with historical pre-pandemic estimates of 75-90% prevalence among computer user (Anshel 2005, Sheedy et al. 2007, Yan et al. 2008, Rahman and Sanip, 2011, Ranasinghe et al., 2016), suggesting that the shift to exclusive online learning may have simply extended occupational CVS patterns to academic populations.

The observed symptom profile offers important insights into the pathophysiology of CVS during prolonged

academic screen use. The higher prevalence of internal symptoms (ISF = 83.3%) compared to external symptoms (ESF = 56.75%) suggests that CVS in academic settings is primarily driven by visual processing demands rather than ocular surface factors alone. Key indicators include 68% of participants reporting extra focusing effort and 68.7% experiencing photophobia, reflecting accommodation and pupillary stress responses. These findings align with existing research, such as a transient myopia shift observed in 20% of VDT workers after prolonged screen exposure due to reduced accommodative power (Luberto et al., 1989). Additionally, studies show a larger lag of accommodation and a 25-30% slower reading speed on digital screens compared to print (Chu et al., 2011), with screen brightness and contrast further affecting performance (Hue et al., 2013). Smaller smartphone screens may also increase accommodative strain due to prolonged reading demands (Bababekova et al., 2011; Kim et al., 2017).

Despite being secondary to accommodative issues (ESF=56.75), dry eye symptoms remain clinically significant, with 78% reporting ocular fatigue, 76.7% requiring lid closure for relief, and 72.7% experiencing eye heaviness. These correlate with reduced blink rates (Patel et al., 1991), increased tear evaporation (Rosenfield, 2011), and a high prevalence of incomplete blinks in VDT users (Portello et al., 2013). Beyond accommodative strain, vergence demands, extraocular muscle fatigue, and tear film abnormalities may also contribute to blurred vision during screen use (Alemayehu & Alemayehu, 2019).

The robust gender difference in CVS severity ($p<0.01$ across all subscales) persists even CVSS17's unidimensionality was validated via Rasch analysis (Gonzalez-Perez et al. 2018), suggesting genuine biological and behavioural mediators. Endocrine related factors may contribute, as estrogen receptors are present in corneal epithelium (Auw-Haedrich & Feltgen, 2003) and progesterone-mediated meibomian gland dysfunction (Versura et al. 2014). Additionally, academic behavioural patterns are more apparent, where female students demonstrate 23% longer continuous study periods (Lenskiy & Paprocki, 2016) and 40% higher likelihood of multi device use (Kim et al. 2017). The 65:35 female: male ratio in our sample mirror national tertiary enrolment trends (Ahmad, 2009), reinforcing the need for gender sensitive interventions. These findings align with prior research demonstrating higher visual symptom scores among female VDT operators in Japan (Shima et al., 1993) and greater visual fatigue in female systems analysts in Brazil (Rocha & Debert-Ribeiro, 2004).

CONCLUSION

This study reveals a striking prevalence of Computer Vision Syndrome (CVS), affecting over 80% of university students engaged in online education during the COVID-19 pandemic. As digital screen use remains integral to modern academia, these findings underscore an urgent need for institutional action. Proactive measures such as structured screen-time breaks, ergonomic training, and awareness programs should be embedded into academic frameworks to alleviate the burden of CVS. Addressing this issue holistically will not only enhance student well-being but also safeguard long-term visual health in an increasingly digital learning landscape.

DECLARATION

The authors declare no competing interests. This research did not receive any specific grant from public, commercial, or not-for-profit funding agencies.

REFERENCES

Alemayehu, A. M., & Alemayehu, M. M. (2019). Pathophysiologic Mechanisms of Computer Vision Syndrome and its Prevention: Review. *World Journal of Ophthalmology & Vision Research*, *2*(4), 1–7.

Aminah Ahmad (2009) Gender Differences and Trends in the Participation of Malaysians in Education: Implications on Employment Outcomes. *International Journal of Knowledge Management Studies* · January 2009

Anshel, J. (Ed.). (2005). *Visual Ergonomics Handbook*. CRC Press. https://doi.org/10.1201/9781420032055

Auw-Haedrich, C., & Feltgen, N. (2003). Estrogen receptor expression in meibomian glands and its correlation with age and dry-eye parameters. *Graefe's Archive for Clinical and Experimental Ophthalmology 241*(9), 705–709. https://doi.org/10.1007/s00417-003-0699-4

Bababekova, Y., Rosenfield, M., Hue, J. E., & Huang, R. R. (2011). Font size and viewing distance of handheld smart phones. *Optometry and Vision Science: Official Publication of the American Academy of Optometry*, *88*(7), 795–797. https://doi.org/10.1097/OPX.0b013e3182198792

Bali J, Navin N, Thakur BR. (2007) Computer vision syndrome: a study of the knowledge, attitudes and

practices in Indian ophthalmologists. *Indian J Ophthalmol*. Jul-Aug;55(4):289-94. https://doi: 10.4103/0301-4738.33042.

Berolo, S., Wells, R. P., & Amick, B. C. (2011). Musculoskeletal symptoms among mobile hand-held device users and their relationship to device use: A preliminary study in a Canadian university population. *Applied Ergonomics*, *42*(2), 371–378. https://doi.org/10.1016/j.apergo.2010.08.010

Chawla, A., Lim, T. C., Shikhare, S. N., Munk, P. L., & Peh, W. C. G. (2019). Computer Vision Syndrome: Darkness Under the Shadow of Light. Can Assoc Radiol J *70*(1), 5–9. https://doi.org/10.1016/j.carj.2018.10.005

Chu, C., Rosenfield, M., Portello, J. K., Benzoni, J. A., & Collier, J. D. (2011). A comparison of symptoms after viewing text on a computer screen and hardcopy. *Ophthalmic Physiol Opt* *31*(1), 29–32. https://doi.org/10.1111/j.1475-1313.2010.00802.x

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Ganne, P., Najeeb, S., Chaitanya, G., Sharma, A., & Krishnappa, N. C. (2021). Digital Eye Strain Epidemic amid COVID-19 Pandemic—A Cross-sectional Survey. *Ophthalmic Epidemiology*, *28*(4), 285–292. https://doi.org/10.1080/09286586.2020.1862243

González-Pérez, M., Susi, R., Barrio, A., & Antona, B. (2018). Five levels of performance and two subscales identified in the computer-vision symptom scale (CVSS17) by Rasch, factor, and discriminant analysis. *PLOS ONE*, *13*(8), e0202173. https://doi.org/10.1371/journal.pone.0202173

Hassan, Z., Szalai, E., Berta, A., Modis, L., & Nemeth, G. (2013). Assessment of Tear Osmolarity and Other Dry Eye Parameters in Post-LASIK Eyes. *Cornea*, *32*(7), e142–e145. https://doi.org/10.1097/ICO.0b013e318290496d

Sitaula, K.R., & Khatri, A. (2018). Knowledge, attitudes and practice of Computer Vision Syndrome among medical students and its impact on ocular morbidity. J Nepal Health Res Counc, *16*(3), 291–296. https://doi.org/10.3126/jnhrc.v16i3.21426

Kim, J., Hwang, Y., Kang, S., Kim, M., Kim, T.-S., Kim, J., Seo, J., Ahn, H., Yoon, S., Yun, J. P., Lee, Y. L., Ham, H., Yu, H. G., & Park, S. K. (2016). Association between Exposure to Smartphones and Ocular Health in Adolescents. *Ophthalmic Epidemiol*, *23*(4), 269–276. https://doi.org/10.3109/09286586.2015.1136652

Lee, H. S., Park, S. W., & Heo, H. (2016). Acute acquired comitant esotropia related to excessive Smartphone use. *BMC Ophthalmology*, *16*, 37. https://doi.org/10.1186/s12886-016-0213-5

Lema, A. K., & Anbesu, E. W. (2022). Computer vision syndrome and its determinants: A systematic review and meta-analysis. *SAGE Open Medicine*, *10*, 20503121221142402. https://doi.org/10.1177/20503121221142402

Lenskiy, A., & Paprocki, R. (2016). *Blink Rate Variability during resting and reading sessions* (arXiv:1605.02037). arXiv. https://doi.org/10.48550/arXiv.1605.02037

Parihar, J. K. S., Jain, V. K., Chaturvedi, P., Kaushik, J., Jain, G., & Parihar, A. K. S. (2016). Computer and visual display terminals (VDT) vision syndrome (CVDTS). *Med J Armed Forces India*, *72*(3), 270–276. https://doi.org/10.1016/j.mjafi.2016.03.016

Patel, S., Henderson, R., Bradley, L., Galloway, B., & Hunter, L. (1991). Effect of visual display unit use on blink rate and tear stability. *Optom Vis Sci. 68*(11), 888–892. https://doi.org/10.1097/00006324-199111000-00010

Portello JK, Rosenfield M, Chu CA. (2013) Blink rate, incomplete blinks and computer vision syndrome. *Optom Vis Sci.* May;90(5):482-7. doi: 10.1097/OPX.0b013e31828f09a7

Rahman, Z. A., & Sanip, S. (2011). Computer User: Demographic and Computer Related Factors that Predispose User to Get Computer Vision Syndrome. *Int J Bus Humanit Technol*, 1(2), 84-91

Ranasinghe, P., Wathurapatha, W. S., Perera, Y. S., Lamabadusuriya, D. A., Kulatunga, S., Jayawardana, N., & Katulanda, P. (2016). Computer vision syndrome among computer office workers in a developing country: An evaluation of prevalence and risk factors. *BMC Research Notes*, *9*, 150. https://doi.org/10.1186/s13104-016-1962-1

Reddy, S. C., Low, C. K., Lim, Y. P., Low, L. L., Mardina, F., & Nursaleha, M. P. (2013). Computer vision syndrome: A study of knowledge and practices in university students. *Nepalese Journal of Ophthalmology*, *5*(2), 161–168. https://doi.org/10.3126/nepjoph.v5i2.8707

Rosenfield, M. (2011). Computer vision syndrome: A review of ocular causes and potential treatments. *Ophthalmic & Physiological Optics: Ophthalmic Physiol Opt* *31*(5), 502–515. https://doi.org/10.1111/j.1475-1313.2011.00834.x

Shantakumari, N., Eldeeb, R., Sreedharan, J., & Gopal, K. (2014). Computer use and vision-related problems among university students in ajman, United arab emirate. *Ann Med Health Sci Res. 4*(2), 258–263.

Sheedy, J. E., Gowrisankaran, S., & Hayes, J. R. (2005). Blink rate decreases with eyelid squint. *Optometry and Vision Science* 82(10), 905–911. https://doi.org/10.1097/01.opx.0000181234.63194.a7

Turgut, B. (2018). Ocular Ergonomics for the Computer Vision Syndrome. *Journal of Eye and Vision*, 1(1).

Versura, P., Giannaccare, G., & Campos, E. C. (2014). Sex-steroid imbalance in females and dry eye. *Current Eye Research*, *40*(2), 162–175. https://doi.org/10.3109/02713683.2014.966847

Wimalasundera, S. (2009). Computer vision syndrome. *Gelle Medical Journal* (1). *11*(1) 25-29 https://doi.org/10.4038/gmj.v11i1.1115

Yan, Z., Hu, L., Chen, H., & Lu, F. (2008). Computer Vision Syndrome: A widely spreading but largely unknown epidemic among computer users. *Computers in Human Behavior*, 24(5), 2026–2042. https://doi.org/10.1016/j.chb.2007.09.004

Yuan, C. W., Hon, E. L. K. H., Lam, O. H., Tay, J., Yue, N. Y., & Viji, K. C. (2022). Computer vision syndrome among university students: A cross sectional study from Perak, Malaysia. *Quest International Journal of Medical and Health Sciences*, *5*(2), Article 2. https://doi.org/10.5281/zenodo.7542787

ORIGINAL ARTICLE

First Words and Expressive Vocabulary Growth in Toddlers: A Cross-Sectional Study

Nur Atiqah Abd Talib[1], Nur Hanisah Tukiran[2,*], W.A Wan Aslynn[2], Marina L. Alisaputri[2]

[1]Excel Qhalif Playgroup & Pusat Terapi Kanak-kanak Kota Damansara, 59-1, Jalan Cecawi 6/33, Kota Damansara, 47810 Petaling Jaya, Selangor
[2]Department of Audiology & Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
first words; vocabulary size; toddlers; word class

ABSTRACT

Background: The early years of life are critical for language acquisition, with toddlerhood marking a pivotal period for vocabulary expansion. Early words offer valuable insight into children's cognitive, linguistic, and social development, yet evidence from Malaysian populations remains limited. This study aims to: (i) identify the first words produced by toddlers aged 12-35 months; (ii) analyse the word classes represented in these first words; (iii) determine the expressive vocabulary size of the toddlers; and (iv) determine the effect of age group on early vocabulary size in toddlers. **Method:** A cross-sectional survey design was employed involving 40 caregivers of typically developing toddlers in Kuantan, Malaysia. Data were collected via a questionnaire comprising demographic information and six items assessing expressive vocabulary, including the first words, the first 10 words, and the estimated total vocabulary size. Descriptive statistics summarised lexical content, and the Mann–Whitney U test compared vocabulary size across age groups. **Result:** Toddlers' first words were predominantly nouns (90%), especially parental terms such as mama and ayah. Analysis of the first 10 words showed a similar dominance of nouns (65.1%), followed by verbs (18.6%) and social terms (5.1%). Vocabulary size differed significantly by age, with younger toddlers clustered in the <10 and 10-49-word ranges, and older toddlers demonstrating broader vocabularies, some exceeding 500 words. **Conclusion:** The findings indicate that Malaysian toddlers' early lexicons mirror cross-linguistic trends, with a strong bias towards nouns. However, vocabulary growth appeared to be slower than established norms, possibly due to the small sample size. These results underscore the need for culturally sensitive benchmarks.

INTRODUCTION

The early years of a child's development are critical for language acquisition, with toddlerhood being particularly pivotal for vocabulary growth. The emergence of first words represents an essential milestone that not only reflects cognitive and linguistic maturation but also signals broader communicative abilities fundamental to social interaction. Mastery of vocabulary has been identified as a cornerstone of language competence, forming the basis for subsequent syntactic and pragmatic development (Kelley & Goldstein, 2019). The expansion of vocabulary in the second year of life reflects the interaction of neurological maturation, environmental input, and growing social-cognitive awareness (Tomasello, 2003).

Numerous studies have shown that children's first words typically emerge toward the end of the first year, followed by a period of gradual vocabulary accumulation and, eventually, a rapid acceleration in word learning often described as the "vocabulary spurt" (Chow & Wehby, 2018; Hung & Chang, 2024). This developmental progression has been consistently documented across various linguistic and cultural contexts, highlighting both universal and language-specific features of early lexical growth (Bornstein et al., 2004).

In a cross-linguistic investigation, Bornstein et al. (2004) examined twenty-month-old children acquiring Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. They found that children with vocabularies ranging from 51 to 200 words produced significantly more nouns than any other word class (verbs, adjectives, and

* *Corresponding author.*
E-mail address: hanisahtukiran@iium.edu.my

closed-class words). However, this noun advantage was not evident among children with fewer than 50 words. Similarly, Hung and Chang (2024), in their longitudinal study of Mandarin-speaking children, reported that children at 20 months produced slightly more nouns than verbs, but verbs became dominant by 26 and 36 months. Although the populations in these two studies differed in age, they are comparable. Typically developing children between 18 and 23 months usually possess 70–100 words (Rescorla et al., 2014), placing 20-month-old children in Hung and Chang's study within the same vocabulary range (51–200 words) as those in Bornstein's study. This suggests that the predominance of nouns during early lexical development is a typical pattern across both Western and East Asian languages, though the timing and rate of shift toward more verb use may vary by language typology and input characteristics.

Despite extensive data from Western and East Asian languages such as English, Spanish, and Mandarin, there remains limited evidence from Malaysia, particularly regarding how early lexical categories and vocabulary size develop during toddlerhood. Understanding these patterns in Malaysian children is crucial for establishing culturally and linguistically appropriate benchmarks for early language assessment and for identifying deviations that may indicate developmental delays. Therefore, this study aims to (i) identify the first words produced by toddlers aged 12–35 months; (ii) analyse the word classes represented in these first words; (iii) determine the spoken vocabulary size of the toddlers; and (iv) examine the effects of age group (12–23 months vs. 24–35 months) on vocabulary size.

METHODOLOGY

Sample Size

An a priori power analysis was conducted using *G*Power 3.1 to estimate the required sample size for detecting age-related differences in vocabulary size between groups. Effect size estimation was informed by Rescorla et al. (2014), who reported a very large effect for age on vocabulary development (partial $\eta^2 = .38$), with older children demonstrating significantly higher vocabulary scores than younger children. This corresponds to a large standardised effect (Cohen's $f \approx 0.78$).

Based on a two-tailed significance level of $\alpha = .05$ and a desired statistical power of 0.80 for detecting a large effect, *G*Power indicated a minimum required total sample of 34 participants for a two-group comparison. To improve reliability and accommodate potential data loss, the target sample size was set at 40 participants. This sample size yields an estimated achieved power exceeding 0.85 for detecting large age-related effects in vocabulary size.

Participants

The inclusion criteria included caregivers of typically developing toddlers aged 12-35 months residing in Kuantan, Malaysia. As Malay and English are the most widely used languages in Malaysia, this study only included caregivers of toddlers who used Malay or English as their primary language. For this study, primary language was operationally defined as the language(s) the toddler most frequently used to communicate at home, as reported by caregivers in the survey. Caregivers were asked to estimate the percentage of exposure to each language when more than one language was used. Following the operational definition proposed by Levey and Cruz (2003), toddlers were classified as bilingual if caregivers reported approximately equal exposure to Malay and English during adult–child interactions (i.e., 50% input in each language). Toddlers were classified as monolingual if a single language accounted for the majority of daily communication.

In addition, the exclusion criteria included a history of neurological conditions, sensory impairment (e.g., hearing loss), autism spectrum disorder, intellectual disability, global developmental delay, or other diagnosed developmental conditions known to affect language acquisition.

Instrumentation

A survey consisting of two components was developed and administered. The first component collected personal information, including the toddler's date of birth, gender, name, medical history (if any), and the caregiver's relationship to the toddler. The second component comprised six questions designed to assess spoken word production. These questions addressed (1) the toddler's primary language(s), (2) the first word produced, (3) expressive ability compared with peers, (4) the first ten words produced, (5) length of utterance currently produced, and (6) estimated vocabulary size. To facilitate responses to the fourth question, participants were provided with examples of early words categorised according to the MacArthur-Bates Communicative Development Inventory (CDI) checklist (Fenson et al., 2007). In addition, for Question 6 (estimated vocabulary size), caregivers were provided with structured categorical ranges (<10, 10–49, 50–199, 200–499, >500 words) rather than open-ended responses to reduce estimation error and improve consistency in reporting. All questions were prepared and administered bilingually in Malay and English to accommodate participants' language

preferences and to maximise comprehension and response accuracy.

Before data collection, the survey was reviewed by the research team, which included a certified speech-language therapist and an academic with expertise in language development, each with more than five years of professional experience. Due to the descriptive and exploratory nature of the study, formal reliability testing was not conducted.The finalised list of questions was then used in both pen-and-paper surveys and an online platform.

Procedures

This cross-sectional study employed convenience sampling and was approved by the International Islamic University Malaysia Research Ethics Committee. In this study, the researcher used the same set of questions for both the online platform and the pen-and-paper format.

The study was advertised through personal messages, viral Facebook postings, and forwarded messages via WhatsApp. For participants who were interested in responding via the online survey, a link to the survey was provided in the advertisement. For those who preferred the pen-and-paper version, the survey was distributed to them personally. Informed consent was obtained from all participants prior to data collection. Specifically, written consent was obtained from participants completing the paper-based questionnaire, while online participants provided electronic consent before accessing the survey. To protect confidentiality, all responses were anonymised using identification codes, and no personally identifying information was retained in the dataset.

All questionnaires were completed independently by the participants. To ensure that participants met the inclusion criteria, reminders of the requirements were provided prior to them beginning the survey. These reminders were included in the viral messages, as well as in the first section of the online survey, before participants provided their consent.

Data Analysis

All data were analysed using IBM SPSS Statistics Version 27. Descriptive statistics, including frequencies and percentages, were computed to summarise the children's vocabulary size and word classes. Prior to inferential analysis, the normality of vocabulary size distributions was assessed using the Shapiro–Wilk test, as the sample size was less than 50. The results indicated a non-normal distribution ($p < 0.05$), warranting the use of a non-parametric approach. Accordingly, the Mann–Whitney U test was employed to compare vocabulary size between the two age groups (12–23 months versus 24–35 months).

RESULTS

A total of 40 caregivers of typically developing toddlers participated in the study. The analysis of caregiver reports indicated that 18 toddlers (45%) were males, while 22 toddlers (55%) were females. In terms of language background, 24 toddlers (60%) were monolingual Malay speakers, three toddlers (7.5%) were monolingual English speakers, and 13 toddlers (32.5%) were bilingual speakers of Malay and English. For analysis purposes, the toddlers were divided into two age groups: 12-23 months (n = 24) and 24-35 months (n = 16). Participant demographics are summarised in Table 1.

Table 1: Demographic characteristics of toddlers (N = 40)

Variables	n	%
Gender		
• Male	18	45.0
• Female	22	55.0
Age Group		
• 12-23 months	24	60.0
• 24-35 months	16	40.0
Language Background		
• Malay monolingual	24	60.0
• English monolingual	3	7.50
• Malay-English bilingual	13	32.5

First Words

Analysis of toddlers' very first words showed that the majority were nouns (n = 36, 90%), followed by verbs (n = 3, 7.5%) and exclamations (n = 1, 2.5%). Within the noun category, the most frequent first words referred to parents. Specifically, 15 toddlers (37.5%) produced words for "father," including "papa", "*ayah", "abah", "baba"*, and "daddy", while 17 toddlers (42.5%) produced words for "mother," such as "*mama", "ibu", "mak"*, and "mommy".

In addition to parental terms, several other words appeared as toddlers' first utterances, reflecting both basic needs and salient elements of the environment. The verb "*nak"* (want) was reported by three toddlers (7.5%), suggesting that expressions of desire emerge early as part of social interactions. Two toddlers (5%) used the word "*air"* (water), a term strongly tied to daily routines such as drinking. Single reports were also noted for *"kereta"* (car) (n = 1, 2.5%) and *"atuk"* (grandfather) (n = 1, 2.5%).

Word Class Composition of the First 10 Words

Caregivers reported a total of 312 words across toddlers' first 10 words. The majority of these words were nouns (n = 203, 65.1%), followed by verbs (n = 58, 18.6%), social terms (n = 16, 5.1%), adjectives (n = 10, 3.2%), pronouns (n = 8, 2.6%), interjections (n = 6, 1.9%), onomatopoeia (n = 6, 1.9%), adverbs (n = 3, 1.0%), and prepositions (n = 2, 0.6%). This distribution indicates that nouns were by far the largest category of words reported in the early lexicons of toddlers in this study.

Within the noun category, words referring to family members and names accounted for the largest proportion (n = 109, 53.7%). Other common subcategories of nouns included food and drinks (n = 28, 13.8%), animals (n = 26, 12.8%), and body parts (n = 19, 9.4%). Less frequently reported nouns included those related to furniture and household items (n = 8, 3.9%), transportation (n = 7, 3.5%), clothing (n = 3, 1.5%), and toys (n = 3, 1.5%).

With regards to verbs, the commonly reported verbs included *"nak"* (want), *"duduk"* (sit), and *"makan"* (eat). Social terms reported included greetings and expressions such as "hi". Adjectives (n = 10, 3.2%) included descriptive words such as "*sakit"* (pain) and *"sejuk"* (cold). Pronouns (n = 8, 2.6%) included items such as "*itu"* (that) and "*ini"* (this).

The most frequently reported words overall were *"nak"* (n = 20, 6.4%), *"mama"* (n = 15, 4.8%), *"susu"* (milk) (n = 11, 3.5%), *"ibu*" (n = 10, 3.2%), *"ayah"* (n = 10, 3.2%), and *"papa"* (n = 10, 3.2%). Other common words included *"abah"* (n = 8, 2.6%), *"abang"* (brother) (n = 8, 2.6%), and *"air"* (water) (n = 8, 2.6%). In addition, words such as "cat", *"kakak"* (sister), and *"tak nak"* (do not want) were each reported by six toddlers (1.9% each). These high-frequency items represent the lexical entries most commonly found within the toddlers' first 10 words, particularly nouns and verbs. A detailed distribution of words across categories is presented in Table 2.

Vocabulary Size

To analyse vocabulary size, toddlers were classified into five groups: category 1, fewer than 10 words; category 2 10–49 words; category 3, 50–199 words; category 4, 200–499 words; and category 5, more than 500 words. The mean vocabulary category for toddlers aged 12–23 months (n = 24) was 1.71 (SD = 0.86), reflecting the clustering in the lower categories. Ten toddlers (41.7%) were reported to produce fewer than 10 words, and 13 toddlers (54.2%) were in the 10–49-word range. None of the toddlers in this age group were reported to have vocabularies of 50 to 199 words or 200 to 499 words. Interestingly, one toddler (4.2%) demonstrated a vocabulary exceeding 500 words, indicating an exceptionally advanced lexical profile compared to peers in the same age range.

The mean vocabulary category for toddlers aged 24-35 months (n = 16) was 3.13 (SD = 1.36), indicating a broader distribution across larger vocabulary ranges. The largest proportion (n = 6, 37.5%) was in the 50-199-word range. One toddler (6.2%) had a vocabulary between 200 and 499 words, while four toddlers (25.0%) exceeded 500 words. Smaller vocabulary sizes were less frequent in this age group: three toddlers (18.8%) were in the 10–49-word range, and only two toddlers (12.5%) produced fewer than 10 words.

Overall, younger toddlers (12–23 months) were clustered in the <10 and 10–49-word ranges, while older toddlers (24–35 months) were more evenly distributed across larger vocabulary ranges, including several with vocabularies exceeding 500 words. The full distribution of vocabulary sizes by age group is illustrated in Figure 1.

Table 2: The word class of the first 10 words reported by caregivers. Numbers in brackets indicate the frequency of reported words.

Word Classes/ Category	Words (in English)	Words (in Malay)
Nouns (203)		
Food and drinks (28)	Milk, banana	*Gula, Susu (11), nasi (4), air (8), mi, 'nenen'*
Animals (26)	Duck (2), cat (6), dinosaur (2), bird, fish (2), dog (3), monkey, Shark	*Burung, ayam, cicak, haiwan, gajah (2), Kuda, kucing*
Transportation (7)	bus, car (2), lorry	*Kereta (2), lori*
Toys (3)	Ball (3)	
Clothing (3)	Shoe (2)	*kain*

Family members/ Names (109)	Papa (9), mama (15), mom, mommy (3), daddy (3)	Parents: *ibu (10), abah (8), mak (2), ayah (10), abi, umi, baba (4),* Grandparents: *Wan, atok (6), opah, nenek (5), nana, tok ki* Siblings' nicknames: *abang (8), kakak (6), adik, acih, usu, yayah, along, baby* Cartoon character: *Bing, Didi* Own's name: *Jah, Wawa*
Body parts (19)	teeth, tummy, eye (2), leg (2), finger, hair (2), foot, head, nose	*tangan (2), mata, mulut, kaki (2), bibir*
Furniture and house items (8)	pillow, chair, TV	*Kerusi, bunga, lampu, kipas, pintu*
Verbs (55)	Stop, brush, shower, want	*Nak (20), tak nak (6), makan (5), main, mandi (2), duduk (5), pergi, bak, boleh, jap, keluar, mintak, jom, duduk, tidur (2), minum, tak payah*
Adjectives (10)		*Sakit (2), habis, geli, diam, Sejuk (2), panas, pedas, lapar*
Pronoun (8)		*Itu (3), apa, ini (3), lagi*
Adverb (3)		*tak (2)*
Preposition (2)		*Tepi, nu (disana)*
Interjection (9)	yeay (2), wow, okay (3)	*wah (3)*
Onomatopoeia (6)	Meow (5), woof	
Social terms (16)		Hi (16)

Differences in Vocabulary Size Across Age Groups

A Mann-Whitney U test was conducted to examine whether there were differences in vocabulary size between the two age groups of toddlers. The analysis revealed a statistically significant difference in vocabulary size between toddlers aged 12-23 months (Mdn = 2.0) and those aged 24-35 months (Mdn = 3.0), U = 69.50, Z = -3.56, $p < 0.001$. The effect size was large (r = 0.56, 95% CI [0.30, 0.74]), indicating that younger toddlers generally produced fewer words and consequently had lower vocabulary size scores than older toddlers.

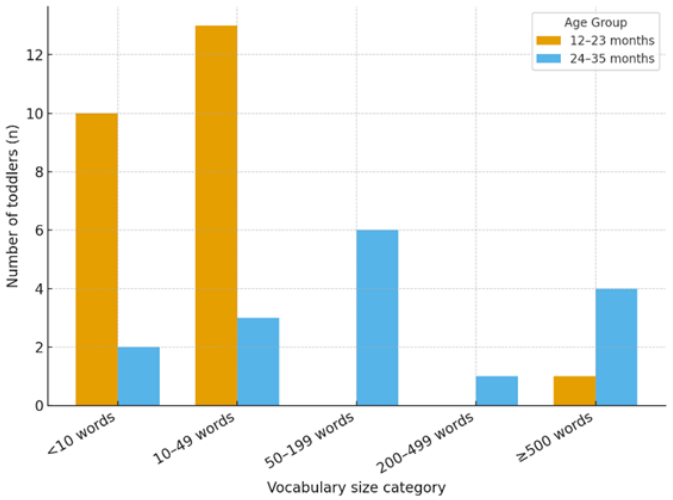


Figure 1: Vocabulary size distribution by age group (12-23 months and 24-35 months).

DISCUSSION

This study examined early lexical development in toddlers by analysing first words, word class composition of the first 10 words, and vocabulary size across two age groups (12–23 months and 24–35 months). The first words were predominantly nouns, especially parental terms such as "mama" and "ayah." Analysis of the word classes of the earliest 10 words revealed the same finding, in which family names, food, and animals were the most frequent, while verbs and other word classes appeared less frequently. Vocabulary size differed significantly between groups: younger toddlers were clustered in the <10 and 10-49-word ranges, whereas older toddlers were more evenly spread across larger ranges, with some exceeding 500 words.

Despite individual variation, the early lexical categories observed in this study were consistent with those reported in other populations. Previous research has shown that children's first words often cluster around important people, social routines, animals, and foods (Schneider et al., 2015; Tardif et al., 2008), and our findings reflected a similar distribution. Family members accounted for the largest proportion of first words, followed by food and drinks, animals, and body parts. Such parallels highlight the universality of early lexical content across languages and cultures, suggesting that children's initial vocabularies are rooted in the immediate and familiar environment. Large-scale cross-linguistic work has further shown that frequently heard words, particularly concrete nouns, are acquired earlier (Braginsky et al., 2019). In line with this, the predominance of parental names and other salient nouns in our data underscores the role of input frequency and communicative relevance in shaping the earliest lexicons. By contrast, Hung and Chang (2024) found that although nouns slightly outnumbered verbs at 20 months, verbs became dominant by 26 and 36 months. Such developmental shifts could not be captured in our study, which relied solely on parent reports of the earliest 10 words.

Vocabulary size further illustrated developmental progression in our sample, although the pace appeared somewhat slower than that reported in other studies. Younger toddlers (12-23 months) were clustered in the <10 and 10-49-word ranges, while the older group (24-35 months) displayed more advanced lexical growth, with the majority exceeding 50 words and some surpassing 500 words. In contrast, normative data suggest that typically developing children between 18 and 23 months usually have 70–100 words (Rescorla et al., 2014), and by 30 months, often reach 150-200 words (Rescorla et al., 2000, 2014).

While our findings confirm the positive relationship between vocabulary size and age, the lower ranges in the younger group and the wide variability in the older group suggest that lexical growth in this sample may appear less advanced.. This difference may be partly attributable to the relatively small sample size, which could magnify variability and limit generalizability. Additionally, culturally specific interaction patterns may have influenced vocabulary growth. Studies conducted in non-Western contexts such as Iran and Turkey have shown that mothers' directives during parent–child interactions tend to be more dominant than other communicative behaviours (Ashtari et al., 2020; Diken, 2009). Directives are more likely to shut down conversation rather than invite children to respond verbally, thereby limiting verbal participation and potentially hindering language development (Ratner, 2013; Conway et al., 2017). Other unmeasured factors, including socioeconomic status and parental education, may also have contributed and should be considered in future research.

Of particular concern, a subset of children in the 24–35-month group were still reported by parents as having fewer than 50 words, and these parents still perceived this level of vocabulary as normal, indicating a limited awareness of developmental milestones. Clinically, this is significant because restricted expressive vocabulary at this age is considered a red flag for language delay. Although the percentile benchmarks proposed by Thal and Tobias (1994) were derived from Western populations, they are commonly used as reference points for identifying children at risk. Children older than 24 months may be considered at risk if their vocabulary size falls below the 10th percentile of a normative sample or if they produce fewer than 50 words or fail to combine words (Buschmann et al., 2009; Rescorla, 1989). We acknowledge that Malaysian-specific norms are not yet established, and caution should be exercised when interpreting these benchmarks. These findings emphasise the importance of parent education and the development of local normative data for early identification of language delays.

CONCLUSION

This study highlights that toddlers' earliest vocabularies are dominated by nouns, particularly parental names and other familiar items, reflecting both universal patterns and the influence of environmental frequency. Vocabulary size increased with age, although growth appeared somewhat slower than normative benchmarks. Some children over 24 months were still reported to have fewer than 50 words, a clinical red flag often overlooked by parents. While the small sample size and reliance on parent-reported first

words limit generalizability, the findings highlight the importance of cross-linguistic perspectives and increased parental awareness in supporting the timely identification of language delays.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Ashtari, A., Samadi, S. A., Yadegari, F., & Harooni, G. G. (2020). The relationship between Iranian maternal verbal responsiveness styles and children's communication acts with expressive and receptive vocabulary in 13–18-month-old typically developing children. *Early Child Development and Care, 190*(15), 2392–2401. https://doi.org/10.1080/03004430.2019.1573227

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., Pêcheux, M.-G., Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-Linguistic Analysis of Vocabulary in Young Children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. Child Development, 75(4), 1115–1139. https://about.jstor.org/terms

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and variability in children's word learning across languages. Open Mind, 3, 52–67. https://doi.org/10.1162/opmi_a_00026

Buschmann, A., Jooss, B., Rupp, A., Feldhusen, F., Pietz, J., & Philippi, H. (2009). Parent based language intervention for 2-year-old children with specific expressive language delay: A randomised controlled trial. Archives of Disease in Childhood, 94(2), 110–116. https://doi.org/10.1136/adc.2008.141572

Chow, J. C., & Wehby, J. H. (2018). Associations Between Language and Problem Behavior: a Systematic Review and Correlational Meta-analysis. Educational Psychology Review, 30(1), 61–82. https://doi.org/10.1007/s10648-016-9385-z

Conway, L. J., Levickis, P. A., Smith, J., Mensah, F., Wake, M., & Reilly, S. (2017). Maternal communicative behaviours and interaction quality as predictors of language development: Findings from a community-based study of slow-to-talk toddlers. *International Journal of Language & Communication Disorders, 52*(6), 717–732. https://doi.org/10.1111/1460-6984.12352

Diken, I. H. (2009). Turkish mothers' self-efficacy beliefs and styles of interactions with their children with language delays. *Early Child Development and Care, 179*(4), 425–436. https://doi.org/10.1080/03004430701200478

Fenson, L., Marchman, V., Thal, D., Dale, P., Reznick, J., & Bates, E. (2007). MacArthur Communicative Development Inventories: Users' Guide and Technical Manual (2nd Edition). Baltimore: Paul H Brookes.

Hung, Y. F., & Chang, C. J. (2024). Developing early lexical composition in Mandarin-speaking children: A longitudinal study. Journal of Child Language, 51(4), 903–924. https://doi.org/10.1017/S0305000922000654

Kelley, E. S., & Goldstein, H. (2019). Examining Performance on a Process-Based Assessment of Word Learning in Relation to Vocabulary Knowledge and Learning in Vocabulary Intervention. Seminars in Speech and Language, 40(5), 344–358. https://doi.org/10.1055/s-0039-1688447

Levey, S., & Cruz, D. (2003). The First Words Produced by Children in Bilingual English/Mandarin Chinese Environments. Communication Disorders Quarterly, 24(3), 129–136. https://doi.org/10.1177/15257401030240030401

Ratner, N. B. (2013). Why talk with children matters: Clinical implications of infant- and child-directed speech research. Seminars in Speech and Language, 34(4), 203–214.

Rescorla, L. (1989). The Language Development Survey: A Screening for Delayed Language in Toddlers. Journal of Speech and Hearing Disorders, 54(November), 587–599.

Rescorla, L., Frigerio, A., Sali, M. E., Spataro, P., & Longobardic, E. (2014). Typical and Delayed Lexical Development in Italian. Journal of Speech, Language, and Hearing Research, 57, 1792–1803. https://doi.org/10.1044/2014

Rescorla, L., Mirak, J., & Singh, L. (2000). Vocabulary growth in late talkers: Lexical development from 2;0 to 3;0. Journal of Child Language, 27(2), 293–311. https://doi.org/10.1017/S030500090000413X

Schneider, R. M., Yurovsky, D., & Frank, M. C. (2015). Large-scale investigations of variability in children's first

words. Proceedings of the Annual Meeting of the Cognitive Science Society, 0, 37.

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's First 10 Words. Developmental Psychology, 44(4), 929–938. https://doi.org/10.1037/0012-1649.44.4.929

Thal, D., & Tobias, S. (1994). Relationships Between Language and Gesture in Normally Developing and Late-Talking Toddlers. Journal of Speech and Hearing Research, 37, 157–170.

Tomasello, M. (2003). Constructing a Language: A Usage-Based Theory of Language Acquisition. Harvard University Press.

ORIGINAL ARTICLE

Vegetable Preferences among Preschoolers in Kuantan, Pahang

Nursyafiqah Shaharudin[1], Pramitha Sari[2], Nurul Hazirah Jaafar[1,3,*]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2]Department of Nutrition, Faculty of Health Sciences, Universitas Alma Ata, Yogyakarta, Indonesia
[3]Children Health and Wellbeing Research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
vegetable preferences; parental feeding practices; preschoolers; children; Malaysia

ABSTRACT

Background: Vegetable intake among Malaysian children remains insufficient, despite its role in supporting healthy growth. While previous studies have explored overall vegetable intake and contributing factors, limited research has specifically focused on children's vegetable preferences and their association with parental feeding practices. Therefore, this study aimed to identify the vegetables preferred by preschoolers, assess common parental feeding practices, and examine the relationship between these practices and children's preferences. **Methods:** A total of 44 caregiver-child pairs were recruited through convenience sampling. Children aged 4 to 6 years and their primary caregivers were included, with those having medical conditions affecting diet excluded. Data were collected using validated questionnaires, including the Comprehensive Feeding Practices Questionnaire (CFPQ) and visual-based vegetable preference assessment. Data were analysed using descriptive statistics to summarise vegetable preferences and parental feeding practices, and Pearson's correlation to assess the relationship between them. **Results:** A total of 44 caregiver-child pairs participated in the study. Carrot, broccoli, and tomato were the most preferred vegetables, while mustard leaf, capsicum, and lady's finger were the least liked. The most common feeding practices reported were "restriction for weight control," "modelling," and "monitoring." A significant negative correlation was observed between the "child control" subscale and vegetable preferences (r = -0.300, n = 44, p = 0.048), while a significant positive correlation was found with "encourage balance and variety" (r = 0.314, n = 44, p = 0.038). **Conclusion:** Specific parental feeding practices, especially those that are supportive and involve structured feeding strategies, particularly those promoting balance and variety, are essential for fostering vegetable acceptance and healthier dietary habits from an early age.

INTRODUCTION

According to the Southeast Asian Nutrition Surveys, vegetable intake among Malaysian children remains below recommended levels, with an average consumption of only 1.07 servings per day, which is considerably lower than the two servings advised in the Malaysian Dietary Guidelines (Chong et al., 2017). Early childhood is a critical period for establishing lifelong eating habits, yet vegetables, despite their richness in vitamins, minerals, and dietary fibre, are often the least accepted food group among young children. Low intake during this formative stage may have long-term implications for diet quality and overall health (Tarro et al., 2022).

Previous studies have shown that children tend to prefer sweeter and milder-tasting vegetables, such as carrots and corn, while bitter-tasting and dark green varieties are less favoured (Cox et al., 2021; Sha An Ali et al., 2020). These preferences are influenced by multiple factors, including sensory properties and familiarity, frequency of exposure, parental influence, as well as cognitive developmental aspects (Estay & Escalona, 2025; Kähkönen, et. al., 2021). As each vegetable contributes distinct micronutrients, limited variety in consumption may result in nutritional gaps, underscoring the need to understand which vegetables are more readily accepted by children.

In shaping children's dietary preferences, positive feeding practices such as modelling healthy eating and maintaining a structured meal environment have been linked to healthier eating habits (Wood et al., 2020). Conversely, excessive child control over food choices or the use of coercive feeding strategies may reduce willingness to try vegetables. International evidence, particularly from

* *Corresponding author.*
E-mail address: hazirahjaafar@iium.edu.my

Western contexts, also indicates that early and repeated exposure, including during breastfeeding and complementary feeding, can enhance acceptance.

Despite this evidence, research in Malaysia examining vegetable preferences among preschoolers and their association with parental feeding practices is limited. Cultural norms, local dietary patterns, and parenting approaches may influence both the vegetables to which children are exposed and the feeding strategies employed by caregivers. Without context-specific evidence, it is challenging to develop effective and culturally appropriate interventions to improve vegetable consumption in young children.

Therefore, this study aimed to determine the vegetables most preferred by preschoolers in Kuantan, assess common parental feeding practices, and examine the relationship between these practices and children's vegetable preferences. The findings are expected to provide valuable insights for caregivers, educators, and policymakers in designing strategies that foster greater acceptance of vegetables and support the establishment of healthy eating habits from an early age.

MATERIALS AND METHODS

Participants

A cross-sectional study was conducted among caregivers and their preschool children aged 4 to 6 years enrolled at IIUM Educare, Kuantan. Recruitment was carried out using convenience sampling, whereby all eligible caregiver-child pairs were invited to participate. Children with serious medical conditions that could affect feeding behaviours or food preferences, including diabetes, cystic fibrosis, cancer, food allergies, and autism spectrum disorder, were excluded.

Comprehensive Feeding Practices Questionnaire

Parental feeding practices were assessed using the validated Malay version of the Comprehensive Feeding Practices Questionnaire (CFPQ) by Shohaimi et al. (2014), which contains 39 items across 12 subscales (Cronbach's alpha = 0.56 to 0.86). Caregivers responded using a five-point scale, with higher scores indicating greater use of the respective feeding practice. Mean scores were calculated for each subscale to compare the prevalence of different practices.

Vegetable Preferences Questionnaire

Children's vegetable preferences were assessed using the adapted questionnaire by Sha An Ali et al. (2020), originally developed by Hanbazaza et al. (2015). The tool consisted of 15 images of locally available vegetables, and children indicated their preference for each item using a binary response ('1 = yes' and '0 = no'). The total 'yes' responses were summed to obtain a vegetable preference score (maximum = 15), with higher scores indicating a broader preference. This tool demonstrated strong internal consistency (Cronbach's alpha = 0.926) and was suitable for use in a care centre setting.

Statistical Analysis

A descriptive analysis was performed to outline participants' characteristics, vegetable preference data, and parental feeding practices. Pearson's correlation test was employed to assess the relationship between parental feeding practices and vegetable preferences. Statistical significance was set at $p < 0.05$, and all analyses were conducted using the Statistical Package for the Social Sciences (SPSS) version 26.0.

Ethical Approval

This study received ethical approval from the International Islamic University Malaysia Research Ethics Committee (IREC) (IIUM/504/14/11/2/IREC 2025-KAHS/DNS).

RESULTS

Characteristics of Participants

In Table 1, 44 caregiver-child pairs participated in the study. Most caregivers were aged between 30 and 39 years (61.4%) and were mothers (79.5%). All respondents were Malay (100%), with the majority holding a degree or higher qualification (72.7%), employed full-time (93.2%), and earning RM10,000 or more per month (29.5%). Most caregivers breastfed their children for six months or longer (61.4%) and introduced solid foods after six months of age (79.5%). The mean age of the children was 4.6 years (SD ± 0.8), with 54.5% being girls.

Table 1. Sociodemographic characteristics of participants (N=44)

Characteristics	Mean (SD)	n (%)
Caregiver characteristics		
Age (years)	37.2 (±4.7)	
21 - 29		2 (4.5%)
30 - 39		27 (61.4%)
40 - 49		14 (31.8%)
50 – 59		1 (2.3%)
Gender		
Male		9 (20.5%)
Female		35 (79.5%)
Race		44 (100%)
Malay		
Relationship status		
Mother		
Father		35 (79.5%)
		9 (20.5%)
Educational level		
Degree/Master/PhD		32 (72.7%)
Diploma		10 (22.7%)
SPM		2 (4.5%)
Employment status		
Employed		41 (93.2%)
Self-employed		1 (2.3%)
Not working		2 (4.5%)
Household income	10,519.51	
≤ RM 1,000	(±6,768.87)	0
RM 1,001 - RM 2,500		0
RM 2,501 - RM 5,000		11 (25.0%)
RM 5,001 - RM 7,500		7 (15.9%)
RM 7,501 - RM 1,0000		10 (22.7%)
≥ RM 10,000		13 (29.5%)
*missing		3 (6.8%)
Breastfeeding duration (months)		
Never		2 (4.5%)
< 1		3 (6.8%)
1 to < 3		6 (13.6%)
3 to < 6		6 (13.6%)
≥ 6		27 (61.4%)
Age of introduction to solid food (months)		
4 to 6		9 (20.5%)
> 6		35 (79.5%)
Children characteristics		
Age (years)	4.7 (±0.8)	
4		23 (52.3%)
5		12 (27.3%)
6		9 (20.5%)
Gender		
Boy		20 (45.5%)
Girl		24 (54.5%)

Preferences of Vegetables among Children

The mean total score for overall vegetable preferences was 8.73 (SD ± 6.04), indicating that children, on average, liked slightly more than half of the vegetables listed. As shown in Figure 1, carrot, broccoli, and tomato were the most preferred, each liked by more than two-thirds of participants, while capsicum and lady's finger were the least preferred, with fewer than half of the children indicating a preference.

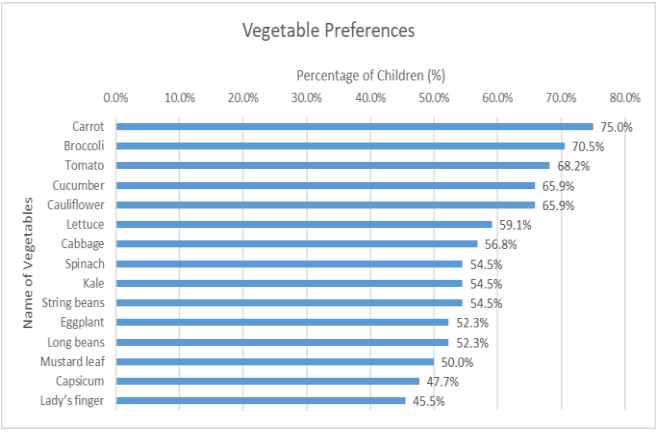


Figure 1. Percentage of vegetable preferences among children (N=44)

Parental Feeding Practices

As shown in Table 2, the highest mean score was recorded for restriction for weight control (19.59 ± 5.95), followed by modelling (16.11 ± 3.15) and monitoring (15.89 ± 3.09). The lowest mean scores were observed for teaching about nutrition (7.77 ± 1.94), food as a reward (6.86 ± 1.42), and emotional regulation (6.16 ± 2.32).

Table 2. Parental feeding practices of caregivers (N=44)

Parental feeding practices (maximum score)	Mean (SD)
Child control (20)	10.73 (±2.25)
Emotion regulation (15)	6.16 (±2.32)
Encouragement of balance and variety (15)	12.86 (±1.75)
Environment (10)	8.39 (±1.43)
Food as reward (10)	6.86 (±1.42)
Involvement (15)	12.14 (±2.29)
Modelling (20)	16.11 (±3.15)
Monitoring (20)	15.89 (±3.09)
Pressure (15)	9.66 (±2.23)
Restriction for health (10)	8.05 (±2.18)
Restriction for weight control (35)	19.59 (±5.95)
Teaching about nutrition (10)	7.77 (±1.94)

Association Between Parental Feeding Practices and Children's Vegetable Preferences

As shown in Table 3, only two parental feeding practice subscales were significantly associated with vegetable preferences. Child control showed a negative correlation with total vegetable preferences (r = -0.300, n = 44, p = 0.048), whereas encouragement of balance and variety showed a positive correlation (r = 0.314, n = 44, p = 0.038).

Table 3: Association between subscales of parental feeding practices and vegetable preferences (N=44)

Parental feeding practices subscales	Vegetable preferences (*r*)	*p*-value
Child control	-0.300	0.048*
Emotion regulation	-0.234	0.127
Encouragement of balance and variety	+0.314	0.038*
Environment	-0.041	0.790
Food as reward	-0.072	0.642
Involvement	-0.135	0.382
Modelling	+0.010	0.948
Monitoring	+0.077	0.620
Pressure	-0.193	0.208
Restriction for health	-0.036	0.816
Restriction for weight control	+0.075	0.628
Teaching about nutrition	-0.257	0.092

Note: r = correlation coefficient; *$p < 0.05$ indicates statistical significance.

DISCUSSION

This study found that preschoolers in Kuantan had a moderate level of vegetable preference, with carrots, broccoli, and tomatoes as the most liked, while lady's finger, capsicum, and mustard leaf were the least preferred. Similar patterns were observed in Sha An Ali et al. (2020), suggesting that certain vegetables remain consistently appealing from early to middle childhood. Early and repeated exposure, particularly during complementary feeding, may explain the preference for carrots and broccoli, as such practices are known to shape children's taste development (Nekitsing & Hetherington, 2022). Interestingly, visual appeal alone did not predict preference, as brightly coloured capsicum was less liked, indicating that taste familiarity, texture, and parental influence may play a stronger role (Nekitsing et al., 2018; Yang et al., 2022).

Among the 12 parental feeding practices assessed, restriction for weight control recorded the highest mean score, possibly reflecting heightened caregiver concern about rising childhood obesity rates in Malaysia (Chua et al., 2024). While restriction may be intended to promote healthy eating, excessive control can reduce willingness to try vegetables (Mahmood et al., 2021). Conversely, modelling and monitoring, which were also common, are considered more effective strategies for improving diet quality, as children often emulate parents' eating behaviours and benefit from repeated exposure to vegetables (Vaughn et al., 2018). Practices such as teaching about nutrition, using food as a reward, and emotional regulation through feeding were less common in this study. The limited use of food as a reward and feeding for emotional regulation can be beneficial in the context of vegetables as coercive strategies are associated with negative dietary behaviours such as increased preference for sugar-rich foods and emotional overeating. (Jansen et al., 2020).

Significant correlations were observed for two feeding practices. Child control was negatively associated with vegetable preferences, aligning with findings that parent-centered feeding practices, including pressure and verbal directives, are associated with reduced vegetable intake (Kristiansen et al., 2017). In contrast, encouragement of balance and variety showed a positive association, reflecting the benefits of authoritative feeding, which combines structure with responsiveness and has been linked to healthier body mass index and dietary intake in children (Shloim et al., 2015). These results emphasise the importance of balanced, supportive guidance over either excessive control or complete autonomy.

The strengths of the current study include its focus on vegetable preferences, a less frequently explored but critical factor in establishing lifelong healthy eating habits, and its examination of parental feeding practices in a specific care centre setting. However, the findings are limited by the small sample size, potential bias from a low response rate, and the use of raw vegetable images, which may not reflect children's preferences for cooked forms. Additionally, the adapted preference tool, though validated for older children, may have been influenced by the developmental stage of the younger participants. The absence of parents during interviews may also have affected responses.

CONCLUSION

This study found that carrots, broccoli, and tomatoes were the most preferred vegetables among preschool children, reflecting common trends in early childhood food choices and providing insight into vegetables that are more readily accepted in care centre settings. Restriction for weight control emerged as the most frequently applied parental feeding practice, followed by modelling and monitoring, whereas teaching about nutrition, offering food as a reward, and using food for emotional regulation were the

least common. The significant associations observed between child control, encouragement of balance and variety, and children's vegetable preferences underscore the important role of parents in shaping eating behaviours. These findings suggest that supportive and structured feeding strategies, particularly those promoting balance and variety, are essential for fostering vegetable acceptance and healthier dietary habits from an early age.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Chong, K., Lee, S., Ng, S., Khouw, I., & Poh, B. (2017). Fruit and Vegetable Intake Patterns and Their Associations with Sociodemographic Characteristics, Anthropometric Status and Nutrient Intake Profiles among Malaysian Children Aged 1–6 Years. *Nutrients*, *9*(8), 723. https://doi.org/10.3390/nu9080723

Chua, K. Y., Chua, K. Y., Chinna, K., Lim, C. L., & Seneviwickrama, M. (2024). Prevalence of childhood overweight and obesity in Malaysia: a systematic review and meta-analysis. *Clinical and Experimental Pediatrics*, 10.3345/cep.2024.00899. https://doi.org/10.3345/cep.2024.00899

Cox, D. N., Baird, D. L., Rebuli, M. A., Hendrie, G. A., & Poelman, A. A. (2021). Sensory characteristics of vegetables consumed by Australian children. *Public Health Nutrition*, *25*(5), 1205–1216. https://doi.org/10.1017/s1368980021000847

Estay, K., & Escalona, V. (2025). Raw or Cooked? Exploring Vegetable Acceptance Among Chilean Children from Different Socioeconomic Backgrounds. Foods, 14. https://doi.org/10.3390/foods14071133

Hanbazaza, M. A., Triador, L., Ball, G. D. C., Farmer, A., Maximova, K., Alexander First Nation, & Willows, N. D. (2015). The Impact of School Gardening on Cree Children's Knowledge and Attitudes toward Vegetables and Fruit. *Canadian Journal of Dietetic Practice and Research*, *76*(3), 133–139. https://doi.org/10.3148/cjdpr-2015-007

Kähkönen, K., Sandell, M., Rönkä, A., Hujo, M., & Nuutinen, O. (2021). Children's Fruit and Vegetable Preferences Are Associated with Their Mothers' and Fathers' Preferences. Foods, 10. https://doi.org/10.3390/foods10020261

Kristiansen, A., Bjelland, M., Himberg-Sundet, A., Lien, N., & Andersen, L. (2017). Associations between sociocultural home environmental factors and vegetable consumption among Norwegian 3–5-year olds: BRA-study. Appetite, 117, 310-320. https://doi.org/10.1016/j.appet.2017.06.031.

Jansen, P. W., Derks, I. P. M., Mou, Y., Rijen, E. H. M., Gaillard, R., Micali, N., Voortman, T., & Hillegers, M. H. J. (2020). Associations of parents' use of food as reward with children's eating behaviour and BMI in a population-based cohort. *Pediatric Obesity*, *15*(11). https://doi.org/10.1111/ijpo.12662

Langer, S. L., Seburg, E., JaKa, M. M., Sherwood, N. E., & Levy, R. L. (2017). Predicting dietary intake among children classified as overweight or at risk for overweight: Independent and interactive effects of parenting practices and styles. *Appetite*, *110*, 72–79. https://doi.org/10.1016/j.appet.2016.12.011

Mahmood, L., Flores-Barrantes, P., Moreno, L. A., Manios, Y., & Gonzalez-Gil, E. M. (2021). The Influence of Parental Dietary Behaviors and Practices on Children's Eating Habits. *Nutrients*, *13*(4). https://doi.org/10.3390/nu13041138

Nekitsing, C., & Hetherington, M. M. (2022). Implementing a "Vegetables First" Approach to Complementary Feeding. *Current Nutrition Reports*, *11*(2), 301–310. https://doi.org/10.1007/s13668-022-00399-z

Nekitsing, C., Hetherington, M. M., & Blundell-Birtill, P. (2018). Developing Healthy Food Preferences in Preschool Children Through Taste Exposure, Sensory Learning, and Nutrition Education. *Current Obesity Reports*, *7*(1), 60–67. https://doi.org/10.1007/s13679-018-0297-8

Sha An Ali, M., Mohd Nazir, N. A., & Abdul Manaf, Z. (2020). Preference, Attitude, Recognition and Knowledge of Fruits and Vegetables Intake Among Malay Children. *Malaysian Journal of Medical Sciences*, *27*(2), 101–111. https://doi.org/10.21315/mjms2020.27.2.11

Shloim, N., Edelson, L. R., Martin, N., & Hetherington, M. M. (2015). Parenting Styles, Feeding Styles, Feeding Practices, and Weight Status in 4–12 Year-Old Children: A Systematic Review of the Literature. *Frontiers in Psychology*, *6*(6). https://doi.org/10.3389/fpsyg.2015.01849

Shohaimi, S., Yoke Wei, W., & Mohd Shariff, Z. (2014). Confirmatory Factor Analysis of the Malay Version Comprehensive Feeding Practices Questionnaire

Tested among Mothers of Primary School Children in Malaysia. *The Scientific World Journal*, *2014*, 1–11. https://doi.org/10.1155/2014/676174

Tarro, S., Lahdenperä, M., Vahtera, J., Pentti, J., & Lagström, H. (2022). Diet quality in preschool children and associations with individual eating behavior and neighborhood socioeconomic disadvantage. The STEPS Study. *Appetite*, *172*, 105950. https://doi.org/10.1016/j.appet.2022.105950

Vaughn, A. E., Martin, C. L., & Ward, D. S. (2018). What matters most - what parents model or what parents eat? *Appetite*, *126*, 102–107. https://doi.org/10.1016/j.appet.2018.03.025

Wood, A. C., Blissett, J. M., Brunstrom, J. M., et al. (2020). Caregiver influences on eating behaviors in young children: A scientific statement from the American Heart Association. *Journal of the American Heart Association,* *9*(10), e014520. https://doi.org/10.1161/JAHA.119.014520

Yang, W. Y., Ong, S. H., Lee, Y. D., Yen, P. L., Lim, K. Y., Naumovski, N., & Jani, R. (2022). Exploration of Malaysian school-children's food preferences: What do we know? *Journal of Tropical Pediatrics,* *68*(5), fmac075. https://doi.org/10.1093/tropej/fmac075

ORIGINAL ARTICLE

Targeting Diabetes Pathways: An *in silico* Study of Bioactive Compounds from *Mitragyna speciosa* Leaf Extract

Taslima Begum[1], Qamar Uddin Ahmed[1*], Mohd Hafiz Arzmi[2], Syed Adnan Ali Shah[3], Mohd Salleh Rofiee[4], Murni Nazira Sarian[5], ABM Helal Uddin[1], Muhammad Muzaffar Ali Khan Khattak[6]

[1]Drug Discovery and Synthetic Chemistry Research Group, Department of Pharmaceutical Chemistry, Kulliyyah of Pharmacy, International Islamic University Malaysia (IIUM), 25200 Kuantan, Pahang Darul Makmur, Malaysia; taslimabegum790@gmail.com (T.B.), quahmed@iium.edu.my (Q.U.A), abmhelal@iium.edu.my (A.B.M.H.U.)

[2]Department of Fundamental Dental and Medical Sciences, Kulliyyah of Dentistry, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia; hafizarzmi@iium.edu.my (M.H.A.)

[3]Faculty of Pharmacy, Atta-ur-Rahman Institute for Natural Products Discovery (AuRIns), Universiti Teknologi MARA, Cawangan Selangor Kampus Puncak Alam, 42300 Bandar Puncak Alam, Selangor, Malaysia; syedadnan@uitm.edu.my (S.A.S.)

[4]Integrative Pharmacogenomics Institute (iPROMISE), Universiti Teknologi MARA, Cawangan Selangor Kampus Puncak Alam, 42300 Bandar Puncak Alam, Selangor, Malaysia; sallehrofiee@uitm.edu.my (M.S.R.)

[5]Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, 43600, Bangi, Selangor, Malaysia; murninazira@ukm.edu.my (M.N.S.)

[6]Food Security and Public Health Nutrition Research Group (Foster), Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, IIUM, 25200 Kuantan, Pahang Darul Makmur, Malaysia

ABSTRACT

Background: The rapid expansion of urbanization has triggered significant lifestyle shifts, most notably an increase in sedentary lifestyles and the widespread consumption of processed, nutrient-poor foods. These factors have been strongly implicated in the rising prevalence of Diabetes Mellitus. Current treatment strategies for diabetes remain heavily reliant on conventional pharmacotherapy, which often carries risks of drug tolerance and adverse effects. In this context, Kratom (*Mitragyna speciosa*), a traditional medicinal plant native to Southeast Asia and historically used to manage various ailments, presents a promising avenue for the development of novel antidiabetic agents. **Methods:** In this study, kratom leaf extract was prepared using methanol-based maceration of kratom leaf powder to obtain a 100% methanolic extract (100%M). The extract was subsequently analyzed using Q-Tof-LCMS, leading to the putative identification of bioactive compounds. These compounds were further subjected to molecular docking against two key target proteins, namely PDB codes 3A4A and 4N8D, to assess their binding affinities. **Results:** The results revealed several promising candidates including apigenin 7-(2''-*E*-*p*-coumaroylglucoside, kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside, 5,6,7,3',4'-pentahydroxy-8-methoxyflavone 7-apioside, luteolin 7-rhamnosyl(1->6)galactoside and 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside), all of which demonstrated stronger binding affinities than both the standard and native ligands. Further evaluation of physicochemical and pharmacokinetic parameters revealed that 4-*p*-coumaroylquinic acid, scopolin, and emmotin A showed instant drug likeness, indicating their potential for direct drug development. Additionally, other compounds with high binding energies may be optimized through structural modifications to enhance their pharmacological profiles, thereby serving as lead candidates for the development of novel antidiabetic therapeutics. **Conclusion:** The current study has identified promising lead compounds for the development of novel antidiabetic agents, offering valuable guidance for future research on kratom-based antidiabetic drug discovery.

Keywords:
Diabetes, kratom, leaf, PDB, maceration, antidiabetic, drug likeness.

INTRODUCTION

Diabetes is a chronic health condition that arises when the body is unable to produce sufficient insulin or cannot effectively utilize the insulin it produces, leading to elevated blood glucose levels (hyperglycaemia).

It is recognized as one of the fastest-growing global health challenges of the 21st century. In 2024, an estimated 589 million people aged 20-79 were living with diabetes worldwide. This number is projected to increase by 45%, reaching nearly 853 million by 2050. In Southeast Asian countries such as Malaysia, Indonesia, the Philippines, and Thailand, the incidence of diabetes is expected to rise by as much as 73% according to the

Corresponding author.
E-mail address: muzaffar@iium.edu.my, quahmed@iium.edu.my

International Diabetes Federation report (2025). These alarming trends highlight the urgent need for effective strategies in prevention, management, and treatment.

Among the different types of diabetes, type 2 diabetes accounts for the majority of cases worldwide (>90%). Although it can often be prevented, delayed, or even reversed in its early stages (International Diabetes Federation report, 2025), it remains a major global health concern. In addition to lifestyle modifications such as adopting a more active routine and healthier dietary choices, treatment with appropriate antidiabetic drugs is also essential. However, conventional synthetic antidiabetic drugs are frequently associated with tolerance, resistance, and adverse side effects (Pollack et al., 2010). This underscores the urgent need to develop novel antidiabetic agents derived from natural sources, particularly traditional medicinal plants, which have long been used by communities to manage various health conditions.

In Southeast Asia, among the diverse range of medicinal plants, kratom (*Mitragyna speciosa*) (Figure 1) stands out as one of the most prominent species with potential for the development of novel antidiabetic agents, owing to its long history of traditional uses (Begum et al., 2025a). Thus, a range of scientific studies have been performed to investigate the antidiabetic properties of kratom leaf using *in vitro*, *in vivo*, and *in silico* models. For instance, Limcharoen et al. (2022) reported the *in vitro* α-glucosidase inhibitory activity of three different kratom extracts, as well as its major alkaloid, mitragynine. Similarly, Zhang et al. (2023) investigated the antidiabetic potential of kratom ethanol extract through both in vitro α-glucosidase inhibition assays and *in vivo* studies in a rat model, further supporting its therapeutic promise.

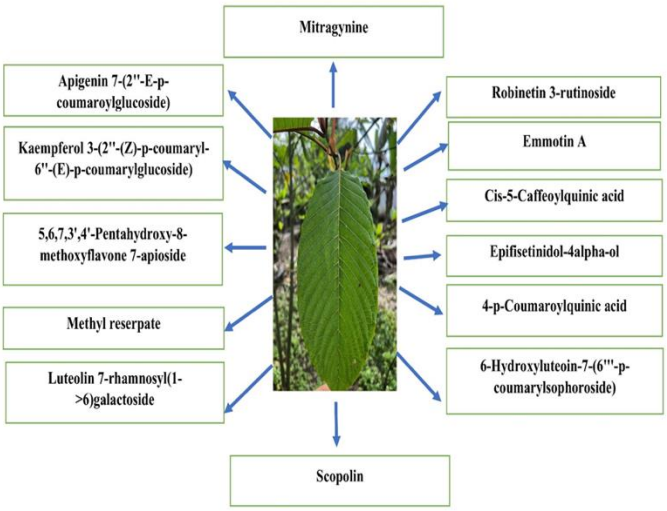


Figure 1: Kratom leaf along with the identified compounds

Furthermore, Hossain et al. (2023) demonstrated the antidiabetic properties of kratom leaves collected from Rangamati, Bangladesh, using both *in vivo* (Swiss albino mice) and *in silico* models (PDB ID: 5T4H). In addition, a clinical study conducted in West Kalimantan, Indonesia, evaluated the blood glucose response to kratom tea in ten participants. The study design involved three phases: administration of standard glucose in the 1st week, followed by administration of standard glucose combined with 1% kratom tea in the 2nd week, and a one-week washout period without treatment. The outcomes exhibited that kratom tea exhibited potential in reducing blood glucose levels (Fadly, 2025).

The objective of this article was to prepare a kratom leaf extract and identify its phytochemical constituents through Q-ToF-LCMS. The identified compounds were subsequently subjected to molecular docking analysis against two key antidiabetic targets: the α-glucosidase enzyme (PDB code: 3A4A) and dipeptidyl peptidase-4 (DPP-4) (PDB code: 4N8D). Finally, physicochemical and pharmacokinetic parameters of the docked compounds were evaluated to assess their potential as lead candidates for the development of novel antidiabetic therapeutics.

MATERIALS AND METHODS

Plant Material Collection and Identification

Kratom leaves were collected from Kedah, Malaysia and taxonomically verified at the Kulliyyah of Pharmacy, International Islamic University Malaysia (voucher no. PIIUM 0358). The leaves were completely cleaned with water to eliminate adhering soil and debris. Later, the cleaned leaves were carefully dried at 40 °C for 72 hours. After that, the dried material was pulverized into a coarse powder using a Universal cutting mill, yielding the powdered kratom leaf material for further analysis.

Extraction Through Maceration

A total of 100 g of dried kratom leaf powder was macerated in 100% methanol at a ratio of 1:3 (w/v) for 24 hours. The mixture was filtered, and the resulting filtrate was concentrated using a rotary evaporator at 40 °C. The recovered methanol was reintroduced to the plant residue, and the extraction process was repeated twice under the same conditions. In total, the procedure was performed three times to obtain 100% methanol maceration extract (100%M) (Begum et al., 2025b).

Profiling of Bioactive Compounds in *M. speciosa* leaf 100%M via Quadrupole Time-of-Flight Liquid Chromatography Mass Spectrometry (Q-ToF-LCMS)

Bioactive compound profiling was performed by employing the Agilent 1290 Infinity and 6550 iFunnel Q-ToF-LCMS (Agilent Technology, Santa Clara, CA, USA) equipped with an electrospray ionization (ESI) source operating in both +ve and -ve ionization modes. For sample preparation, 1 mg of the 100%M extract was allowed to dissolve in 250 µL of LC-grade methanol, followed by vortexing and ultrasonication. Subsequently, 250 µL of distilled water was added, yielding a final volume of 500 µL. The solution was allowed to centrifuge for fifteen minutes and filtered to get a clear solution. Later, it was transferred into a vial for analysis. Chromatographic separation was carried out on a Phenomenex Kinetex C18 column (250 mm × 4.6 mm, 5 µm) maintained at 27°C. The mobile phase consisted of methanol and 0.1% formic acid, with a gradient elution starting at 5% methanol and increasing to 100% methanol. The total run time was 10 minutes with a flow rate of 0.7 mL/min. Mass spectrometric data were acquired in the m/z range of 50 to 1500 with a collision energy ramp of 35 eV. Source parameters were maintained as follows: gas temperature 200 °C, flow rate 14 L/min, nebulizer pressure 35 psig, sheath gas temperature 350 °C, and sheath gas flow 11 L/min. The acquired data were processed using ACD/Spec Manager and pre-processed with MZmine software, including baseline correction, peak detection, filtering, alignment, smoothing, and gap filling. The processed data were exported in CSV format for further analysis (Ananda et al., 2025).

***In silico* Molecular Docking Analysis**

Molecular docking was done to determine the predicted binding affinities of the identified bioactive phytoconstituents at the active sites of the α-glucosidase enzyme (PDB code: 3A4A) and DPP-4 (PDB code: 4N8D) (https://www.rcsb.org/). The three-dimensional (3D) structures of the positive controls, i.e., acarbose for α-glucosidase and sitagliptin for DPP-4, as well as the identified compounds from Q-ToF-LCMS analysis, were retrieved from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/). Docking was performed using AutoDock Vina, Autodock Tools 1.5.7, and Discovery Studio (Trott & Olson, 2010). For α-glucosidase (3A4A), redocking of the native ligand was conducted with grid dimensions (X, Y, Z) set to 26, 26, and 26, respectively, and grid center coordinates (X, Y, Z) set to 21.243, -7.756, and 24.341. For DPP-4 (4N8D), the grid dimensions were set to 26.0, 24.0, and 26.0, respectively, with grid center coordinates (X, Y, Z) of 18.3, 30.424, and 53.897, which were used for redocking of the native ligand. These docking setups ensured that the RMSD values of the redocked native ligands were less than 2 Å, confirming the reliability of the docking protocol. The binding interactions between the proteins and the docked compounds were subsequently analyzed in both two-dimensional (2D) and three-dimensional (3D) representations using Discovery Studio (Mia et al., 2022).

Physico-Chemical and Pharmacokinetic Properties

To predict a drug's behaviour in the body, optimize its therapeutic efficacy, and ensure safety, it is essential to evaluate its physicochemical and pharmacokinetic parameters. Hence, the bioactive compounds identified from the 100%M extract by Q-ToF-LCMS analysis were also investigated for their physicochemical and pharmacokinetic parameters. This evaluation was performed to forecast their drug-likeness and ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) profiles, thereby providing insights into their potential as lead candidates for antidiabetic drug development (http://www.swissadme.ch/index.php, https://biosig.lab.uq.edu.au/pkcsm/prediction) (Ananda et al., 2025).

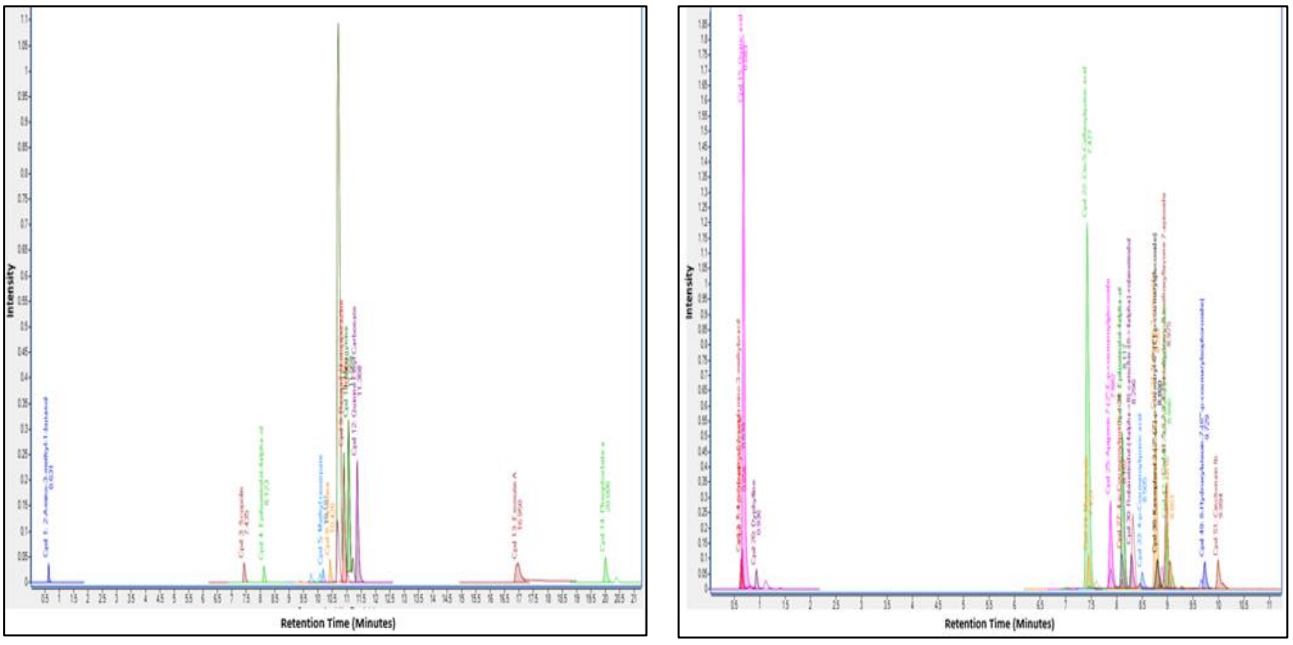
RESULTS

% Yield and Q-ToF-LCMS Analysis of *M. speciosa* leaf 100%M Extract

The 100%M extract (yield: 30.17 %) was found to contain a diverse range of phytochemical constituents, as identified through Q-ToF-LCMS analysis in both positive and negative ionization modes (Table 1). The representative chromatograms obtained from each ionization mode are presented in Figure 2, while the chemical structures of selected identified phytoconstituents are illustrated in Figure 3.

Table 1: Identified compounds of *M. speciosa* leaf 100%M extract

Identified compounds	RT	Score	Formula	m/z
cis-5-Caffeoylquinic acid	7.43	95.67	$C_{16}H_{18}O_9$	353
Scopolin	7.44	99.58	$C_{16}H_{18}O_9$	355
Apigenin 7-(2''-*E*-*p*-coumaroylglucoside)	7.88	97.7	$C_{30}H_{26}O_{12}$	577
4-*p*-Coumaroylquinic acid	8.11	99.75	$C_{16}H_{18}O_8$	337
Epifisetinidol-4alpha-ol	8.12	98.73	$C_{15}H_{14}O_6$	289
Robinetin 3-rutinoside	8.77	95.09	$C_{27}H_{30}O_{16}$	609
Kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside)	8.80	98.64	$C_{39}H_{32}O_{15}$	739
5,6,7,3',4'-Pentahydroxy-8-methoxyflavone 7-apioside	8.98	96.62	$C_{21}H_{20}O_{12}$	463
Luteolin 7-rhamnosyl(1->6)galactoside	8.99	99.07	$C_{27}H_{30}O_{15}$	593
6-Hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside)	9.73	99.52	$C_{36}H_{36}O_{19}$	771
Methyl reserpate	10.18	99.25	$C_{23}H_{30}N_2O_5$	415
Mitragynine	11.07	99.6	$C_{23}H_{30}N_2O_4$	399
Emmotin A	16.95	94.76	$C_{16}H_{22}O_4$	279



A B

Figure 2: Q-ToF-LCMS chromatograms of both positive (A) and negative (B) ionization modes of *M. speciosa* leaf 100%M extract

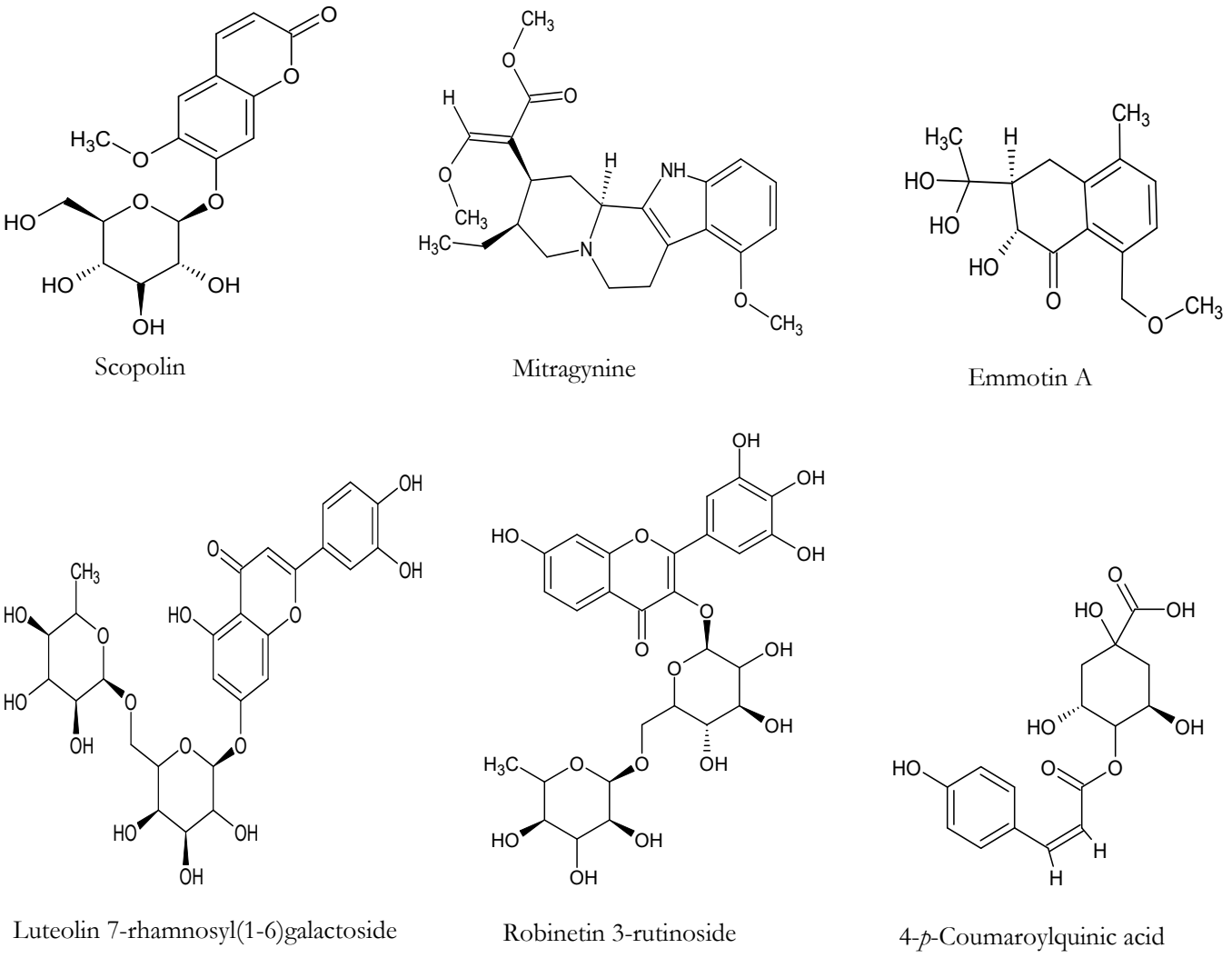


Figure 3: Chemical structures of some of the identified compounds of *M. speciosa* leaf 100%M extract

In silico Molecular Docking Analysis

The putative compounds identified through Q-ToF-LCMS analysis were subjected to molecular docking against two proteins (PDB ID: 3A4A and 4N8D), along with their respective native ligands and standard inhibitors. For PDB ID: 3A4A, the native ligand and the standard drug acarbose showed binding energies of -5.9 kcal/mol and -7.9 kcal/mol, respectively. The control docking of the native ligand yielded an RMSD value of 0.7, confirming the reliability of the docking protocol. Among the 13 compounds analyzed (Table 2), 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside) showed the strongest binding affinity (-11.9 kcal/mol), followed by apigenin 7-(2''-*E*-*p*-coumaroylglucoside) (-11.6 kcal/mol), luteolin 7-rhamnosyl(1-6)galactoside (-10.9 kcal/mol), robinetin 3-rutinoside (-10.0 kcal/mol), and 5,6,7,3',4'-pentahydroxy-8-methoxyflavone 7-apioside (-9.6 kcal/mol). The weakest binding affinity was observed for emmotin A (-7.2 kcal/mol); however, this value was still stronger than that of the native ligand and comparable to the standard acarbose. The binding interactions of 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside), apigenin 7-(2''-*E*-*p*-coumaroylglucoside), luteolin 7-rhamnosyl(1-6)galactoside, scopolin, and emmotin A, along with the native ligand and standard, against PDB ID: 3A4A are illustrated in Figure 4.

Table 2. The binding interactions along with the bond lengths of identified compounds of *M. speciosa* leaf 100%M extract against 3A4A

SI NO	Ligands	Binding affinity (kcal/mol)	Molecular Interactions	
			Hydrogen bond with bond distance (Å)	**Extra interactions with bond distance (Å)**
1.	*cis*-5-Caffeoylquinic acid	-8.4	SER 241; 2.47; ASP 242: 1.88; HIS 2.80: 2.36; THR 310: 2.91; SER 311: 2.30; PRO312: 2.09	TYR 158: 4.10; LYS 156: 5.34
2.	Scopolin	-8.5	SER 157: 2.30	HIS 280: 4.54; TYR 158: 3.73
3.	Apigenin 7-(2''-*E*-*p*-coumaroylglucoside)	-11.6	THR 310: 3.11; ASN 415: 2.24; HIS 112: 2.30; ASP 215: 3.09	VAL 216: 5.03; TYR 72: 5.79; PHE 178: 4.05; ARG 442: 4.31; ASP 307: 4.01; ARG 315: 5.39; HIS 280: 5.61
4.	4-*p*-Coumaroylquinic acid	-8.8	SER 241: 2.12; ASP 242: 2.09	TYR 158: 3.82; ARG 315: 3.78
5.	Epifisetinidol-4alpha-ol	-8.7	TYR 158: 2.47; ASP 215: 2.53; GLU 277: 2.16	VAL 216: 5.50; ASP 352: 4.57; ARG 442: 3.58; ARG 315: 5.35
6.	Robinetin 3-rutinoside	-10.0	ASN 415: 2.67; LEU 313: 2.34; PRO 312: 2.63 & 2.35; ASP 242; 2.52	-
7.	Kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside)	-9.1	ASP 215: 2.91; GLU 411: 2.74 & 2.88; TYR 158: 2.51; LYS 156: 2.69; THR 310: 3.04; SER 311; 2.98; ARG 315: 2.32 & 3.08; ASP 242; 2.33	ASP 307: 4.65 & 3.38; ASP 242: 3.40; PRO 312: 3.89; TYR 158: 5.14; ASP 352; 4.46; VAL 216: 5.17; ASP 352; 4.46
8.	5,6,7,3',4'-Pentahydroxy-8-methoxyflavone 7-apioside	-9.6	ARG 442: 2.45; GLU 277: 2.22; HIS 280: 2.57; ASP 307: 2.65; PRO 312: 2.43; ARG 315: 2.23; LYS 156: 2.92	LEU 313: 3.44; ARG 315: 5.13 & 3.89
9.	Luteolin 7-rhamnosyl(1->6)galactoside	-10.9	PRO 312: 3.20; HIS 280: 3.00; THR 310: 1.78 & 3.02; ASP 307: 3.30 & 3.40; ARG 315: 2.71 & 2.30; ASP 69: 3.24; ASP 215: 3.05	ASP 215: 4.20; VAL 216: 5.09; GLU 277: 4.55; TYR 72: 5.72; PHE 303: 5.29 & 5.23
10.	6-Hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside)	-11.9	ASP 69: 3.24; ASP 215: 3.05; ASP 307: 3.40 & 3.30; ARG 315: 2.71; PRO 312: 3.20; THR 310: 1.78 & 3.02; HIS 280: 3.09	TYR 72: 5.72; ASP 215: 4.20; VAL 216: 5.09; GLU 277: 4.55; PHE 303: 5.29 & 5.23; SER 240: 3.24
11.	Methyl reserpate	-9.4	PRO 312: 2.33; ARG 442: 2.48	PHE 178: 4.62; GLU 411: 4.47 & 4.73; TYR 158: 5.10
12.	Mitragynine	-8.0	HIS 280: 2.63; ARG 442: 2.83	PHE 303: 4.46 & 5.58
13.	Emmotin A	-7.2	ARG 315: 2.45	ARG 315: 4.28
Controls	Native Ligand	-5.9	GLU 277: 2.41 & 2.85; ASP 215: 2.62 & 3.10; HIS 351: 1.95; ARG 442: 1.85; ASP 69: 1.86	-
	Standard (acarbose)	-7.9	ASP 352: 2.79; GLU 277: 2.32; GLU 411: 2.66; HIS 280: 2.41; PRO 312: 1.74; TYR 158; 2.45	-

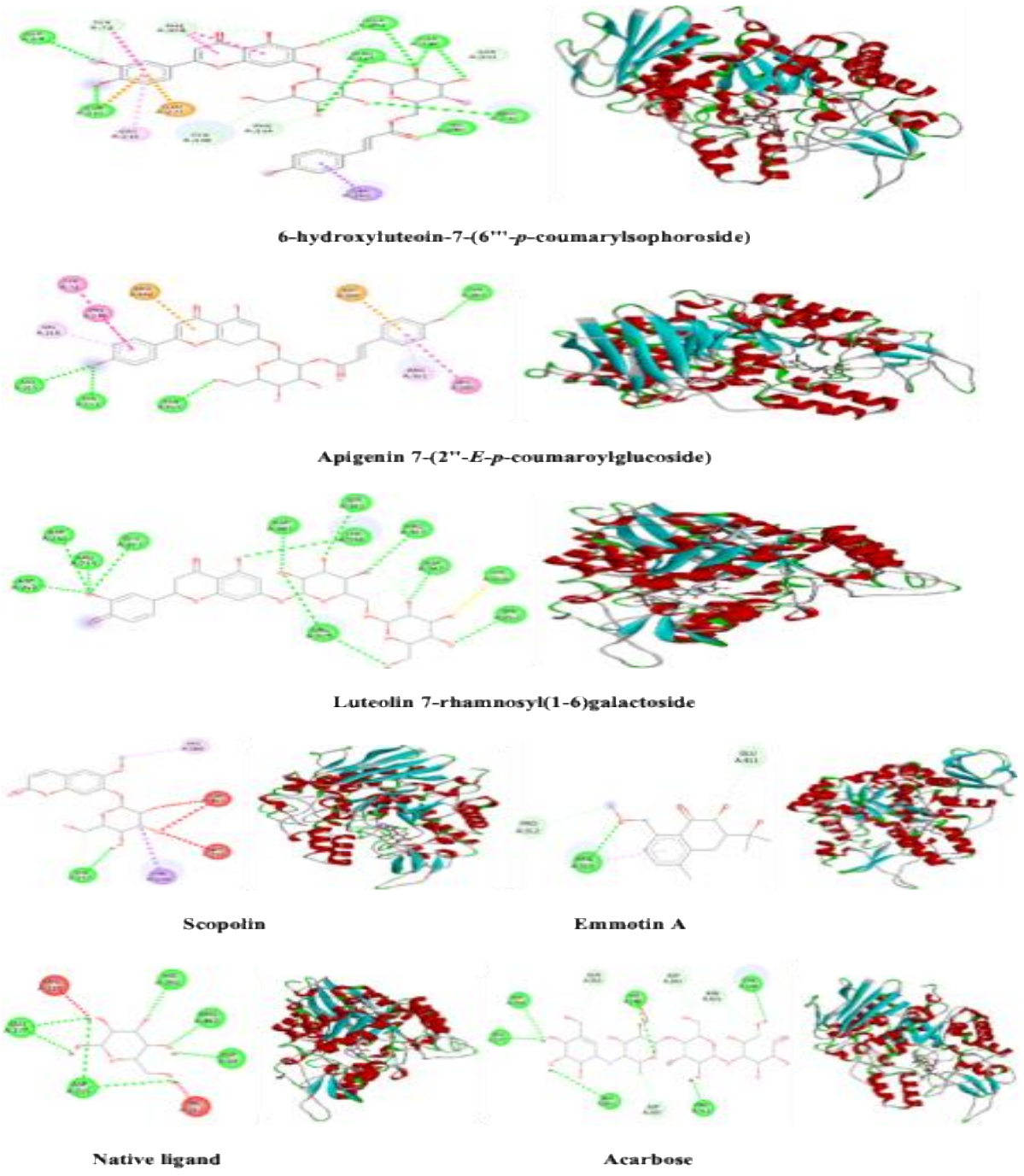


Figure 4: The binding interactions 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside), apigenin 7-(2''-*E*-*p*-coumaroylglucoside), luteolin 7-rhamnosyl(1-6)galactoside, scopolin, and emmotin A, along with the native ligand and standard against PDB ID: 3A4A

The compounds, along with the native ligand and standard inhibitor, exhibited diverse binding interactions with protein 3A4A. The native ligand interacted with the protein through conventional hydrogen bonds involving amino acid residues GLU 277, ASP 215, HIS 351, ARG 442, PRO 312, and ASP 69. For 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside), multiple amino acids contributed to hydrogen bonding and other non-covalent interactions. Specifically, ASP 69, ASP 215, ASP 307, ARG 315, PRO 312, THR 310, and HIS 280 were involved in conventional

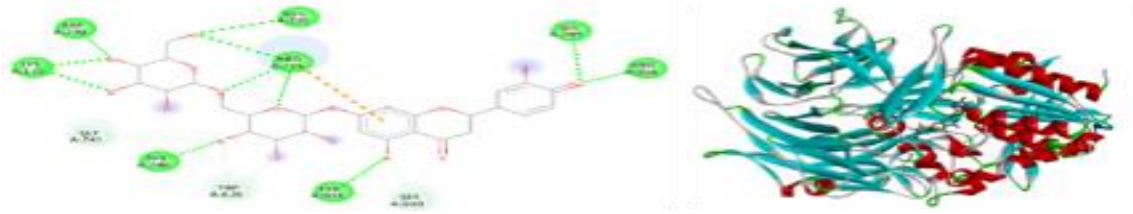
hydrogen bonds. In the case of apigenin 7-(2''-*E*-*p*-coumaroylglucoside), hydrogen bonds were formed with THR 310, ASN 415, HIS 112, and ASP 215. Emmotin A, which showed the lowest binding energy among the analyzed compounds, interacted with ARG 315 through both hydrogen bonding and additional non-covalent interactions. The standard inhibitor (acarbose) formed hydrogen bonds with ASP 352, GLU 277, GLU 411, HIS 280, PRO 312, and TYR 158.

Conversely, for PDB ID: 4N8D, the native ligand and standard inhibitor sitagliptin showed binding energies of -8.4 kcal/mol and -9.4 kcal/mol, respectively. The control docking of the native ligand produced an RMSD value of 1.57, validating the docking protocol. Among the 13 compounds analyzed (Table 3), luteolin 7-rhamnosyl(1-6)galactoside demonstrated the strongest binding affinity (-11.2 kcal/mol), followed by 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside) (-11.0 kcal/mol), kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside) (-10.8 kcal/mol), robinetin 3-rutinoside (-10.3 kcal/mol), and apigenin 7-(2''-*E*-*p*-coumaroylglucoside) (-10.1 kcal/mol). The weakest binding affinity was observed for emmotin A (-7.9 kcal/mol), although this value was still stronger than that of the native ligand. The binding interactions of luteolin 7-rhamnosyl(1-6)galactoside, 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside), kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside), scopolin, and emmotin A, along with the native ligand and standard, against PDB ID: 4N8D are shown in Figure 5.

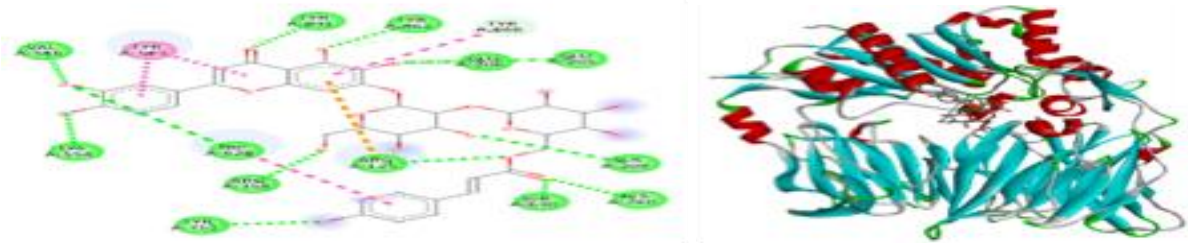
Table 3. The binding interactions, along with the bond lengths of identified compounds of *M. speciosa* leaf 100%M extract against 4N8D

SI NO	Ligands	Binding affinity(kcal/mol)	Molecular Interactions	
			Hydrogen bond with bond distance (Å)	**Extra bond interactions with bond distance (Å)**
1.	*cis*-5-Caffeoylquinic acid	-8.4	SER 630: 2.25; GLU 205: 2.11; HIS 740: 2.58; TRP 629: 2.26; VAL 546: 2.28	-
2.	Scopolin	-8.2	GLU 206: 2.88; ARG 125: 2.34; GLU 205: 2.80	GLU 205: 4.60; TYR 666: 4.99 & 5.64; TYR 662: 4.63
3.	Apigenin 7-(2''-*E*-*p*-coumaroylglucoside)	-10.1	VAL 546: 3.26; ASP 545: 3.35; LYS 554: 2.37; SER 630: 3.02; TYR 631: 2.68; ARG 358: 2.70 & 2.86	TYR 666: 4.11; TYR 547: 4.58 & 4.29
4.	4-*p*-Coumaroylquinic acid	-8.5	ARG 125: 2.93	TYR 547: 4.31
5.	Epifisetinidol-4alpha-ol	-8.1	LYS 554: 2.15; TRP 629: 2.48	TYR 547: 4.18; TYR 666: 5.82
6.	Robinetin 3-rutinoside	-10.3	VAL 546: 2.67; LYS 554: 1.91 & 2.60; TRP 629: 2.12; ARG 125: 2.54; SER 630: 2.47; HIS 740: 2.83;	PHE 357: 4.42 & 5.72; TYR 666: 5.10; TYR 662: 5.07

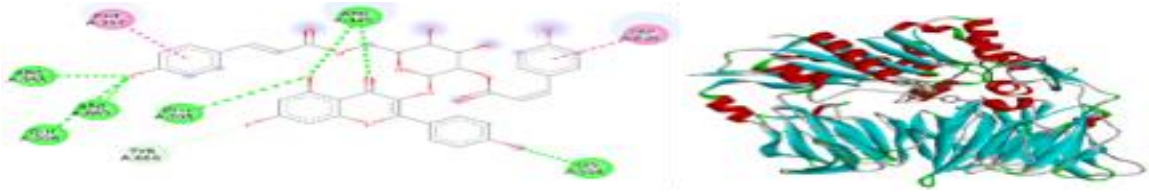
			ASN 710: 2.28; GLU 206: 2.77; ARG 358: 2.04	
7.	Kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside)	-10.8	ARG 358: 2.58; GLU 206: 3.25; ARG 669: 2.84; GLU 205: 3.20; ARG 125: 2.62 & 2.52; LYS 554: 2.11	TRP 629: 3.81; PHE 357: 3.99
8.	5,6,7,3',4'-Pentahydroxy-8-methoxyflavone 7-apioside	-9.2	VAL 546: 2.86; TRP 629: 2.86; GLU 205: 2.14; GLU 206: 2.38; ARG 358: 2.42 & 2.35; SER 209: 2.02	PHE 357: 3.74; TYR 666: 5.48; ARG 125: 4.70; TYR 547: 4.82 & 3.69
9.	Luteolin 7-rhamnosyl(1->6)galactoside	-11.2	SER 209: 2.72; ARG 358: 1.82; TYR 631: 2.28; HIS 740: 3.37; LYS 122: 3.05 & 2.04; ASP 739: 3.30; GLU 205: 2.99; ARG 125: 2.87, 2.55 & 2.71	ARG 125: 4.62
10.	6-Hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside)	-11.0	TYR 752: 2.83; ARG 358: 2.74; ARG 125: 2.20; SER 209: 2.53; HIS 740: 3.10; SER 630: 2.06; GLU 205: 3.59; GLU 206: 3.00; TYR 662: 3.28; TYR 631: 2.15; VAL 546: 3.37; LYS 554: 2.30; TRP 629: 2.91	TYR 666: 5.86; TRP 629: 3.87; ARG 125: 4.45
11.	Methyl reserpate	-8.6	ARG 358: 2.57; GLU 206: 1.95	-
12.	Mitragynine	-8.0	TYR 547: 2.99; TRP 629: 2.78; TYR 631: 2.52; TRP 629: 2.78	TYR 547: 3.88 & 4.41; LYS 554: 5.28
13.	Emmotin A	-7.9	SER 630: 1.99; TRP 629: 2.51	TRP 629: 3.99
Controls	Native Ligand	-8.4	TYR 662: 1.79	PHE 357: 4.50; SER 630: 4.96; TYR 666: 4.96; TYR 662: 4.44; GLU 206: 2.90; GLU 205: 2.03; ASP 663: 4.84
	Standard (sitagliptin)	-9.4	TYR 662: 2.79; HIS 740: 2.58	TYR 662: 4.04; VAL 656: 5.02; TYR 666: 4.66; LYS 554: 4.81; TRP 629: 3.65; ASP 545: 3.32



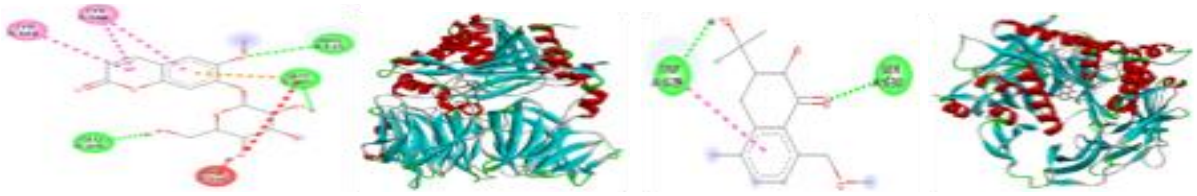
Luteolin 7-rhamnosyl(1-6)galactoside



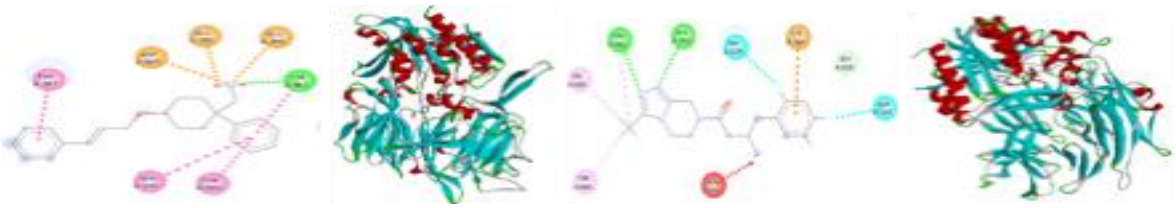
6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside)



Kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside)



Scopolin Emmotin A



Native ligand Sitagliptin

Figure 5: The binding interactions of luteolin 7-rhamnosyl(1-6)galactoside, 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside), kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside), scopolin, and emmotin A, along with the native ligand and standard against PDB ID: 4N8D

Against the protein target 4N8D, the compounds, along with the native ligand and standard inhibitor, exhibited distinct binding interactions. The native ligand formed a hydrogen bond with amino acid TYR 662, and additional non-covalent interactions with amino acid residues PHE 357, SER 630, TYR 666, TYR 662, GLU 206, GLU 205, and ASP 663. For luteolin 7-rhamnosyl(1->6)galactoside, hydrogen bonds were established with SER 209, ARG 358, TYR 631, HIS 740, LYS 122, ASP 739, GLU 205, and ARG 125. Scopolin (-8.2 kcal/mol) interacted through hydrogen bonds with GLU 206, ARG 125, and GLU 205. Emmotin A (-7.9 kcal/mol), which showed the weakest binding affinity among the analyzed compounds, formed hydrogen bonds with SER 630, and TRP 629. The standard inhibitor, sitagliptin exhibited hydrogen bonding interactions with TYR 662, and HIS 740, consistent with its

strong binding affinity.

Physico-Chemical and Pharmacokinetics Properties

All thirteen compounds identified from the *M. speciosa* leaf 100%M extract through Q-ToF LCMS analysis were further evaluated for their physicochemical and pharmacokinetic profiles to understand their drug likenesses and toxicity profiles (Table 5).

Table 5. The physicochemical and pharmacokinetic parameters of compounds identified from *M. speciosa* leaf 100%M extract

Compound	Physicochemical						Pharmacokinetic & Toxicity				
	HA	HD	MW	Log P	V	DL	IA	CR	HT	AM	ROAT
Epifisetinidol-4alpha-ol	6	5	290	-0.02	0	Yes	65.56	0.003	No	Yes	2.45
6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside)	19	11	773	-3.64	3	No	17.24	-0.44	No	No	2.49
Luteolin 7-rhamnosyl(1-6)galactoside	15	9	595	-3.43	3	No	27.78	0.25	No	Yes	2.73
Mitragynine	5	1	399	2.02	0	Yes	92.97	0.86	Yes	No	3.16
Robinetin 3-rutinoside	16	10	611	-3.89	3	No	20.60	-0.38	No	No	2.50
1,4-di-o-caffeoylquinic acid	12	7	517	-0.35	3	No	16.36	-0.06	No	No	2.54
5,6,7,3',4'-pentahydroxy-8-methoxyflavone 7-apioside	12	7	464	-2.32	2	No	41.38	0.44	No	No	2.61
Apigenin 7-(2''-*E*-*p*-coumaroylglucoside	12	6	579	-0.56	3	No	47.64	0.04	No	No	2.62
Emmotin A	4	2	278	0.96	0	Yes	95.02	1.05	No	No	2.04
Methyl reserpate	6	2	415	1.30	0	Yes	93.52	48.12	Yes	No	2.93
Scopolin	9	4	354	-1.23	0	Yes	48.12	0.72	No	No	2.39
Kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside)	15	7	741	3.6953	3	No	63.18	-0.38	No	No	2.51
4-*p*-coumaroylquinic acid	7	5	338	-0.35	0	yes	27.58	0.44	No	No	1.77

Note: HA: hydrogen bond acceptor; HD: hydrogen bond donor; MW: molecular weight, gram/mole; Log P: predicted octanol/water partition coefficient; DL: drug likeness; V: number of violations to Lipinski's rule; IA: intestinal absorption, %; CR: clearance; %; HT: hepatotoxicity; AM: AMES toxicity; ROAT: rat oral acute toxicity, mol/kg

DISCUSSION

The 100%M extract derived from *M. speciosa* leaves using the maceration extraction technique yielded a substantial amount of 30.17%, indicating efficient extraction of bioactive compounds from the source material. This relatively high yield suggests that the extraction method employed was suitable for recovering a broad spectrum of phytochemicals, which is essential for downstream pharmacological or nutraceutical applications (Ahmed et al., 2025).

The Q-ToF-LCMS analysis revealed a diverse array of phytochemical constituents, identified through both positive and negative ionization modes. This dual-mode approach enhances the sensitivity and coverage of compound detection, as certain phytoconstituents ionize more effectively in one mode than the other. Among the identified compounds, several phenolic and flavonoid compounds were detected, including *cis*-5-caffeoylquinic acid, 4-*p*-coumaroylquinic acid, epifisetinidol-4alpha-ol, robinetin 3-rutinoside, kaempferol 3-(2''-(*Z*)-*p*-coumaryl-6''-(*E*)-*p*-coumarylglucoside), 5,6,7,3',4'-pentahydroxy-8-methoxyflavone 7-apioside, luteolin 7-rhamnosyl(1->6)galactoside, 6-hydroxyluteoin-7-(6'''-*p*-coumarylsophoroside), and apigenin 7-(2''-*E*-*p*-coumaroylglucoside). Earlier works have revealed that flavonoids are responsible for various pharmacological activities, including anti-diabetic effects. For instance, rutin has been shown to retard carbohydrate absorption by obstructing the α-glucosidase enzyme activity, kaempferol can increase glucose uptake, and luteolin is capable of inhibiting lipid synthesis (Praparatana et al., 2022). Moreover, 4-*p*-coumaroylquinic acid has been reported to exert antidiabetic activity in albino Wistar rats (Amalan et al., 2016). These findings suggest that the phenolic and flavonoid compounds identified in kratom leaf extract may contribute to its antidiabetic potential. Beyond these, mitragynine has already been documented to possess antidiabetic properties (Limcharoen et al., 2022). Furthermore**,** scopolin (scopoletin-7-O-β-*D*-glucopyranoside) is a glycoside of scopoletin, has also been linked to antidiabetic activity **(**Jang et al., 2020**).**

The tentative phytoconstituents detected via Q-ToF-LCMS analysis were subjected to the molecular docking against two diabetes linked proteins (PDB ID: 3A4A and 4N8D), along with their respective native ligands and standard inhibitors. Compared to the native ligands and standard inhibitors, the compounds identified from the 100%M extract showed strong binding affinities against both diabetes-associated protein targets. Alpha-glucosidase inhibitors, such as acarbose are the drugs that retard carbohydrate absorption from the small intestine by competitively inhibiting the enzymes accountable for converting intricate carbohydrates into simple carbohydrates which are easily absorbed. By delaying carbohydrate absorption, these drugs effectively reduce postprandial blood glucose concentrations (Akmal et al., 2024). Several of the identified compounds showed binding affinities and interaction profiles that were not only comparable to, but in some cases stronger than, acarbose against the alpha-glucosidase enzyme (3A4A). These findings provide a promising baseline for the development of novel alpha-glucosidase inhibitors.

Similarly, DPP-4 inhibitors such as sitagliptin exert their antidiabetic effects through modulation of incretin hormones. Incretins are secreted within minutes of food intake but are rapidly degraded by the DPP-4 enzyme. By inhibiting DPP-4, these drugs prolong incretin activity, thereby enhancing insulin secretion and reducing postprandial and fasting hyperglycemia. Beyond their antihyperglycemic effects, DPP-4 inhibitors have also been reported to exert antihypertensive, anti-inflammatory, antiapoptotic, and immunomodulatory effects on the cardiovascular and renal systems, independent of the incretin pathway (Kasina, & Baradhi, 2019). The identified compounds from the 100%M extract displayed binding affinities and interactions that were comparable to, and in some cases stronger than, those of the native ligand and sitagliptin, suggesting their potential as promising scaffolds for the development of new DPP-4 inhibitors.

The assessment of physicochemical and pharmacokinetic parameters is essential to evaluate the potential of bioactive compounds as lead candidates for the development of new drug products. Accordingly, all thirteen compounds identified from the 100%M extract were further evaluated for their physicochemical and pharmacokinetic profiles. The assessment revealed that out of the 13 compounds, epifisetinidol-4alpha-ol, mitragynine, emmotin A, methyl reserpate, scopolin, and 4-*p*-coumaroylquinic acid showed drug likeness. Among these, epifisetinidol-4alpha-ol, mitragynine, and methyl reserpate were predicted to possess either AMES mutagenicity or hepatotoxicity. In contrast, 4-*p*-coumaroylquinic acid, scopolin, and emmotin A showed no predicted toxicity, with intestinal absorption values of low (27.577%), medium (48.119%) and high (95.021%), respectively. Advances in drug design and formulation strategies may help overcome limitations such as poor absorption or toxicity. Approaches including particle size reduction, increased surface area, and prodrug formation have been successfully employed to enhance bioavailability and reduce adverse effects (Alagga et al.,

2025). Although the other flavonoid compounds showed strong binding affinities against both protein targets, their high molecular weight and hydrophilic nature limited their drug-likeness. Structural optimization, such as the introduction of small hydrophobic groups, could improve their pharmacokinetic properties and enhance their potential as drug candidates (Sarian et al., 2017).

CONCLUSION

The global incidence of diabetes is escalating at an alarming pace, creating an urgent demand for novel antidiabetic therapies that can overcome the limitations, tolerance, and side effects associated with existing drugs. Natural products remain a valuable reservoir for drug discovery, and in this context, kratom leaves represent a promising source of bioactive compounds. The present study establishes a baseline for future research aimed at isolating and developing antidiabetic agents from kratom leaves. Among the identified compounds, 4-*p*-coumaroylquinic acid, scopolin, and emmotin A emerged as particularly noteworthy, demonstrating favourable binding affinities, pharmacokinetic properties, and safety profiles. These findings highlight their potential as lead compounds for the development of novel antidiabetic drugs. Overall, the findings support the potential of the 100%M extract as a rich source of bioactive compounds. Future studies should focus on isolating individual constituents, evaluating their pharmacological effects, and exploring synergistic interactions that may enhance therapeutic efficacy.

ACKNOWLEDGEMENT

This study was partially supported by the CRIGS UKM-IIUM-UiTM-UPM Research Collaboration Grant 2024 (C-RIGS24-017-0023), under the Research Management Centre, International Islamic University Malaysia (IIUM). The authors gratefully acknowledge the financial support provided by Regal Malay Capital Berhad, Malaysia (SPP22-137-0137), and the CPS Sejahtera Award, IIUM. We also extend our sincere appreciation to the Kulliyyah of Pharmacy, IIUM, for providing the research facilities necessary to complete this work.

AUTHORS CONTRIBUTIONS

Conceptualization, Q.U.A.; M.M.A.K.K.; T.B.; methodology, T.B.; Q.U.A.; resources, Q.U.A.; M.H.A.; data curation, T.B.; M.S.R.; S.A.A.S.; writing-original draft preparation, T.B.; writing-review and editing, Q.U.A.; M.M.A.K.K.; M.N.S.; visualization, Q.U.A.; A.B.M.H.U.; M.N.S.; supervision, Q.U.A. All authors have read and agreed to the published version of the manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY

Data will be made available on reasonable request.

FUNDING

No direct funding was received to assist with the preparation of this manuscript.

ETHICAL APPROVAL

Not applicable.

REFERENCES

Ahmed, Q. U., Atirah Ali, S. N. A., Abdul Jalil, N. N. S., Uddin, A. H., Rofiee, M. S., Ali Shah, S. A., Salleh, M. Z., Abbas, S. A., Ahmed, M. Z., Sarker, M. Z. I. & Mia, M. A. R. (2025). Unveiling the anti-diabetic effects of *Garcinia hombroniana* fruit pericarp and leaf: *in vitro* radical scavenging, alpha-glucosidase inhibition, pharmacokinetics, and *in silico* molecular docking of Q-TOF-LCMS identified compounds. *Natural Product Research*, 1-12. https://doi.org/10.1080/14786419.2025.2534861

Akmal, M., Patel, P., & Wadhwa, R. (2024). *Alpha glucosidase inhibitors.* StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK557848/

Alagga, A. A., Pellegrini, M. V., & Gupta, V. (2024). *Drug Absorption.* StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK557405/.

Amalan, V., Vijayakumar, N., Indumathi, D., & Ramakrishnan, A. (2016). Antidiabetic and antihyperlipidemic activity of p-coumaric acid in diabetic rats, role of pancreatic GLUT 2: *In vivo* approach. *Biomedicine & Pharmacotherapy*, *84*, 230-236. https://doi.org/10.1016/j.biopha.2016.09.039

Ananda, Z. K, Begum, T., Ahda, M., Rofiee, M. S., Shah, S. A. A., Salleh, M. Z., Almutairi, B.O., Azmi, S. N. H., Sah, P., Wardani, A. K., Khatib, A., Abbas, S. A., Mia, M. A. R., Ahmed, Q. U. (2025). Comparative evaluation of phytochemical screening, *in vitro* antioxidant & alpha-glucosidase inhibitory properties of *Ceiba pentandra* & *Basella rubra* leaf extracts: Identification of active principles by Q-TOFLCMS, ADMET prediction &

molecular docking approach. *Journal of King Saud University - Science*, *37*(1), 162024. https://doi.org/10.25259/JKSUS_16_2024

Begum, T., Arzmi, M. H., Khatib, A., Uddin, A. H., Aisyah Abdullah, M., Rullah, K., So'ad, S. Z. M., Haspi, N. F. Z., Sarian, M. N., Parveen, H., Mukhtar, S., Ahmed, Q. U. (2025a). A review on *Mitragyna speciosa* (Rubiaceae) as a prominent medicinal plant based on ethnobotany, phytochemistry and pharmacological activities. *Natural Product Res*earch, *39*, 1636-1652. https://doi.org/10.1080/14786419.2024.2371564

Begum, T., Ahmed, Q. U., Arzmi, M. H., Uddin, A. H., Khatib, A., Mohamad, S. N. A. S., Nasarudin, S. N. I. S., Abbas, S. A., Abd Aziz, Z. H. (2025b). Phytochemical screening and evaluation of *in vitro* antioxidant, antibacterial and antibiofilm activities of *Mitragyna speciosa* leaf extract. *Malaysian Journal of Chemistry*, *27*(5), 80-94.

Fadly, D. (2025). *Blood glucose response to kratom tea (Mitragyna speciosa Korth.).* BIO Web of Conferences (Vol. 153, p. 03015). EDP Sciences. https://doi.org/10.1051/bioconf/202515303015

Hossain, M. U., Hossain, M. S., Dey, S., Naimur Rahman, A. B. Z., Chowdhury, Z. M., Bhattacharjee, A., Ahammad, I., Islam, M. N., Moniruzzaman, M., Hosen, M. B., Ahmed, I., Das, K. C., Keya, C. A.,& Salimullah, M. (2023). *In vivo* and *in silico* experiment on *Mitragyna speciosa* offers new insight of antidiabetic potentials. *bioRxiv*, *2023*, 11. https://doi.org/10.1101/2023.11.08.566341

International Diabetes Federation. (2025). *IDF Diabetes Atlas* (11th ed.). IDF. https://diabetesatlas.org/resources/idf-diabetes-atlas-2025/

Jang, J. H., Park, J. E., & Han, J. S. (2020). Scopoletin increases glucose uptake through activation of PI3K and AMPK signaling pathway and improves insulin sensitivity in 3T3-L1 cells. *Nutrition Research*, *74*, 52-61. https://doi.org/10.1016/j.nutres.2019.12.003

Kasina, S. V. S. K., & Baradhi, K. M. (2019). *Dipeptidyl peptidase iv (DPP IV) inhibitors.* StatPearls. https://www.ncbi.nlm.nih.gov/books/NBK542331/

Limcharoen, T., Pouyfung, P., Ngamdokmai, N., Prasopthum, A., Ahmad, A. R., Wisdawati, W., Prugsakij, W., & Warinhomhoun, S. (2022). Inhibition of α glucosidase and pancreatic lipase properties of *Mitragyna speciosa* (Korth.) Havil. (kratom) leaves. *Nutrients*, *14*(19), 3909. https://doi.org/10.3390/nu14193909

Mia, M. A. R., Ahmed, Q. U., Ferdosh, S., Helaluddin, A. B. M., Awal, M. S., Sarian, M. N., Sarker, M. Z. I., & Zakaria, Z. A. (2022). *In vitro, in silico* and network pharmacology mechanistic approach to investigate the α-glucosidase inhibitors identified by Q-ToF-LCMS from *Phaleria macrocarpa* fruit subcritical CO2 extract. *Metabolites*, *12*(12), 1267. https://doi.org/10.3390/metabo12121267

Pollack, M. F., Purayidathil, F. W., Bolge, S. C., & Williams, S. A. (2010). Patient-reported tolerability issues with oral antidiabetic agents: associations with adherence; treatment satisfaction and health-related quality of life. *Diabetes research and clinical practice*, *87*(2), 204-210. https://doi.org/ 10.1016/j.diabres.2009.11.023

Praparatana, R., Maliyam, P., Barrows, L. R., & Puttarak, P. (2022). Flavonoids and Phenols, the Potential Anti-Diabetic Compounds from *Bauhinia strychnifolia* Craib. Stem. *Molecules*, *27*(8), 2393. https://doi.org/10.3390/molecules27082393

Sarian, M. N., Ahmed, Q. U., Mat So'ad S. Z., Alhassan, A. M., Murugesu, S., Perumal, V., Perumal, V., Mohamad, S. N. A. S., Khatib, A., & Latip, J. (2017). Antioxidant and antidiabetic effects of flavonoids: A structure-activity relationship based study. *BioMed research International*, *2017*(1), 8386065. https://doi.org/10.1155/2017/8386065

Trott, O., & Olson, A. J. (2010) AutoDockVina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, *31*, 455-61. https://doi.org/10.1002/jcc.21334

Zhang, P., Wei, W., Zhang, X., Wen, C., Ovatlarnporn, C., & Olatunji, O. J. (2023). Antidiabetic and antioxidant activities of *Mitragyna speciosa* (kratom) leaf extract in type 2 diabetic rats. *Biomedicine & Pharmacotherapy*, *162*, 114689. https://doi.org/10.1016/j.biopha.2023.114689

ORIGINAL ARTICLE

Association Between Field of Study, General Nutrition Knowledge, and Healthy Eating Practice Among IIUM Students

Alea Ameera Akmal binti Kama[1], Noraishah binti Mohammed Nor[1,2*] , Muhammad Ibrahim Muhammad[1,2] & Muhammad Ashraf Rostam[1,2]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
nutrition knowledge; healthy eating practices; field of study; university students

ABSTRACT

Background: Adequate nutrition knowledge and healthy eating practices are essential for promoting overall health, particularly among university students who often encounter lifestyle changes and academic pressure that may compromise their dietary habits. The field of study has an impact on students' nutritional knowledge and eating practices, as the nutrition-related courses are often linked to increased awareness regarding healthy eating. Therefore, this study aimed to investigate the association between the field of study, general nutrition knowledge, and healthy eating practices among students at the International Islamic University Malaysia (IIUM). **Methods:** A cross-sectional study was conducted among 379 undergraduate students from IIUM, recruited using convenience sampling. Data were collected through a self-administered questionnaire consisting of the Malay version of the General Nutrition Knowledge Questionnaire (GNKQ) and the Starting the Conversation: Diet (STC) tool. Statistical analyses, including One-way ANOVA and Kruskal-Wallis H tests, were performed using SPSS Version 29.0. **Results:** Most participants (64.9%, n = 246) demonstrated good nutrition knowledge, particularly students from dietetics, while only 0.5% (n = 2) exhibited poor knowledge. The average eating practice score indicated a moderately unhealthy dietary pattern, with a mean score of 6.75 out of 14. Interestingly, students from non-health courses reported healthier eating behaviours compared to those from health-related fields. Significant associations were observed between the field of study and general nutrition knowledge ($p < 0.001$) and between the field of study and healthy eating practices ($p = 0.016$). **Conclusion:** The findings suggest that the nutritional knowledge among IIUM students is generally good. Students in the dietetics field reported the healthiest eating practices compared to students in other fields of study. These results highlight the need for interventions that not only provide nutrition education but also promote overall behaviour change strategies.

INTRODUCTION

Nutrition knowledge refers to the understanding of dietary guidelines and nutrient sources (Bhawra et al., 2023). It enables individuals to make informed food choices and build a solid foundation for practising healthy eating, while a lack of knowledge may lead to undernutrition or overnutrition, contributing to an abnormal body mass index (BMI) and increased risk of chronic diseases. Moschonis & Trakman, 2023). Previous studies have shown that higher nutrition knowledge is associated with healthier practices (Ong et al., 2021; Serban et al., 2024).

However, studies have shown that nutrition knowledge and eating practices vary significantly depending on one's educational background, particularly the field of study. For instance, students in health-related disciplines often demonstrate stronger abilities in evaluating health information, which is associated with healthier eating habits (Habiba & Koli, 2024; Yandutkina et al., 2024)

Young adults commonly obtain nutrition information from online platforms, traditional media like television and radio, health professionals, family, and peers (Jusoh et al., 2021; Pavlović et al., 2023; Quaidoo et al., 2018). However, information from these sources is not always accurate, making health professionals the most reliable source of information. Despite wide access to nutrition information, many university students still demonstrate poor levels of nutrition knowledge (Tay et al., 2023).

Healthy eating practices involve the adequate intake of

[1*] *Corresponding author.*
E-mail address: ishah@iium.edu.com

both macronutrients and micronutrients to support bodily functions (Cena & Calder, 2020). However, many students report poor dietary behaviours, such as frequent consumption of processed food and sugary drinks. These behaviours are influenced by convenience, affordability, peer influence and academic stress.

The field of study also plays a role in shaping eating practices. Students in health-related disciplines tend to have higher health literacy and better nutrition awareness due to greater exposure to health science (Sukys et al., 2017). However, evidence on whether this advantage translates into healthier eating practices remains inconsistent.

Therefore, this study aims to investigate the association between field of study, general nutrition knowledge, and eating practices among IIUM students. The results can serve as guidance in strengthening campus-based health initiatives, such as developing targeted awareness campaigns to promote healthier food choices for all students.

MATERIALS AND METHODS

Study Design

This research employed a cross-sectional quantitative study design to investigate the impact of different fields of study on general nutrition knowledge and healthy eating practices among university students. Participants were recruited through convenience sampling, a practical method that utilises both online platforms and physical questionnaire distribution to recruit participants.

Study Population

The study population consists of students aged 18 to 26 years old currently enrolled in the IIUM Gombak Campus, Kuantan Campus, and Pagoh Campus. The inclusion criteria include undergraduate IIUM students aged between 18 and 26 years old.

Data Collection

Validated questionnaires were used to collect relevant information from participants. The questionnaire consists of four parts, which were made available in both online and paper formats.

Part 1 contains the sociodemographic data, including age, gender, campus, Kulliyyah, field of study, access to nutrition education and medical issues. Part 2 is the anthropometric data section, including participants' body weight (in kilograms) and height.

Part 3 of the questionnaire covers the revised Malaysian version of the General Nutrition Knowledge Questionnaire, which was modified and validated by Manan et al. (2023) to assess participants' level of nutrition knowledge. This Malaysian version of GNKQ was adopted from Kliemann et al (2016), who revised the original GNK. It consists of four sections, each evaluating different areas of nutrition knowledge. The first section focuses on knowledge of dietary advice from experts, while the second section tests the ability to classify foods based on their nutrient content. The third evaluates selection skills, requiring participants to choose healthier options, and the fourth examines knowledge of the relationship between diet and health. This questionnaire categorises general nutrition knowledge into four levels based on scoring: poor knowledge (0-23), average knowledge (24-46), good knowledge (47-69), and excellent knowledge (70-93).

Part 4 of the survey contains Starting The Conversation (STC): Diet questionnaire. It is a validated eight-item simplified food frequency instrument that assesses eating practices among individuals, covering both the healthy and unhealthy food choices of the general population. This questionnaire employs a scoring system that categorises responses into three columns based on healthfulness, with lower scores indicating the healthiest dietary practices (Mahoney et al., 2024; Paxton et al., 2011).

Statistical Analysis

The data collected was analysed using SPSS version 29.0. The sociodemographic and anthropometry data obtained were analysed using descriptive statistics. One-way ANOVA was used to determine the association between different fields of study and general nutrition knowledge. The Kruskal-Wallis H test was used to identify the association between different fields of study and healthy eating practices.

RESULTS

Characteristics of Participants

A total of 379 IIUM undergraduate students participated in this study (Table 1). The mean age was 21.31±1.35 years, with most respondents falling between the ages of 21 and 23 years (58.0%, n=220), followed by those aged 18-20 years (35.9%, n=136) and 24-26 years (6.1%, n=23). The sample was predominantly female (88.7%, n = 336), while males represented 11.3% (n = 43).

Participants were recruited from three IIUM campuses, including Kuantan (67.5%, n = 256), Gombak (21.1%, n =

80), and Pagoh (11.3%, n = 43). The majority were from the Kulliyyah of Allied Health Sciences (KAHS), followed by the Kulliyyah of Sustainable Tourism and Contemporary Languages (KSTCL) and the Kulliyyah of Pharmacy. By field of study, 44.6% (n=169) of students were enrolled in health-related programmes, while 19.0% (n=72) were dietetics students and 36.4% (n=138) in non-health fields. Furthermore, 78.9% (n = 299) reported having access to nutrition education, while 21.1% (n = 80) did not.

Lastly, regarding the distribution of BMI, nearly half (45.6%, n = 173) of the respondents had a normal BMI. This was followed by students who were underweight at 22.4% (n = 85), overweight at 21.6% (n = 82), and obese at 10.3% (n = 39). These findings show that, while the majority of the students had normal BMI, a significant proportion (31.9%) fell within the overweight and obese categories, which may raise concerns regarding potential health risks. BMI values among respondents ranged from 12.5 kg/m² to 41.9 kg/m², with a mean BMI of 21.93 ± 4.42, which falls within the normal range according to the CPG, Management of Obesity BMI classifications.

Table 1: Demographic Data of Participants

Sociodemographic Characteristics	Categories	Frequency (n)	Percentage	Mean ± SD
Age	18-20	136	35.9	21.31±1.35
	21-23	220	58.0	
	24-26	23	6.1	
	Total	379	100	
Gender	Male	43	11.3	
	Female	336	88.7	
	Total	376	100	
Campus	Gombak	80	21.1	
	Kuantan	256	67.5	
	Pagoh	43	11.3	
	Total	379	100	
Kulliyyah	KICT	9	2.4	
	KAHS	134	35.4	
	KAED	5	1.3	
	KOD	22	5.8	
	KENMS	10	2.6	
	KOED	8	2.1	
	KOE	8	2.1	
	KIRKHS	35	9.2	
	AIKOL	5	1.3	
	KOM	31	8.2	
	KON	7	1.8	
	KOP	39	10.3	
	KOS	23	6.1	
	KSTCL	43	11.3	
	Total	379	100	
Field of Study	Health	169	44.6	
	Health (Dietetics)	72	19.0	
	Non-health	138	36.4	
	Total	379	100	
Access to Nutrition Education	Yes	299	78.9	
	No	80	21.1	
	Total	379	100	
BMI	Obese	39	10.3	21.93±4.42
	Overweight	82	21.6	
	Normal	173	45.6	
	Underweight	85	22.4	
	Total	379	100	

General Nutrition Knowledge of IIUM Students

The general nutrition knowledge levels among the respondents were categorised into four levels: poor, average, good, and excellent. Based on Table 2, most IIUM students (64.9%, n = 246) demonstrated a good level of general nutrition knowledge. Additionally, 26.6% (n = 101) of the respondents demonstrated an excellent level of knowledge, while 7.9% (n = 30) showed an average level. Only a small fraction of students (0.5%, n = 2) fell into the category of poor general nutrition knowledge. The minimum score obtained was 14 marks, and the highest score was 84, with a mean score of 61.97 ± 11.33.

Table 2: General Nutrition Knowledge Level of Participants

General Nutrition Knowledge Level	Frequency (n)	Percentage (%)	Mean ± SD
Poor	2	0.5	61.97±11.33
Average	30	7.9	
Good	246	64.9	
Excellent	101	26.6	
Total	379	100	

Eating Practices of IIUM Students

Healthy eating practices were analyse using the Starting the Conversation (STC): Diet Questionnaire. The results ranged from a minimum of 0.00 to a maximum of 14.00 scores, with a mean score of 6.75 ± 1.96 among 379 respondents. Since the scoring system of the STC Questionnaire indicates that lower scores reflect healthier eating practices, the mean scores suggest that, on average, IIUM students exhibited moderately unhealthy eating practices.

Association between Field of Study and General Nutrition Knowledge among IIUM Students

Table 3 shows a statistically significant difference in General Nutrition Knowledge scores between the three groups, $p < 0.001$. Students from the Dietetics group had the highest mean rank (303.11), followed by students in other health programs (185.30) and non-health students (136.74). The outcome indicates that the field of study is significantly associated with levels of general nutrition knowledge.

Table 3: Association between General Nutrition Knowledge Scores and Field of Study

	N	Mean Rank	df	p-value
Health	169	185.30		
Health (Dietetics)	72	303.11	2	<0.001
Non-health	138	136.74		

Association between Field of Study and Eating Practice Among IIUM Students

Table 4 indicates that dietetics students had the lowest mean score (6.22 ± 1.87), followed by students from the non-health group (6.70 ± 1.97) and health (non-dietetics) students (7.01 ± 1.94), with a p-value of 0.016 ($p < 0.05$) and the F value (2,376) = 4.18. Therefore, the analysis revealed a statistically significant difference in eating practice scores between health, dietetics and non-health students.

Post hoc analyses using the Bonferroni method (Table 5) revealed that the mean difference between the Health (Dietetics) and Health (non-dietetics) groups was statistically significant ($p = 0.013$), with Dietetics students exhibiting significantly healthier eating practices. However, no significant differences were found between Dietetics and Non-health students ($p = 0.267$), or between Health and Non-health students ($p = 0.522$).

These findings suggest that the field of study, particularly specialised education in nutrition and dietetics, may be associated with healthier eating behaviours among young adults.

Table 4: Comparing Eating Practices between Health, Health (Dietetics), and Non-health (n=380)

Variable	N	Healthy Eating Practices		F-statistics (df)	p-value
Field of Study		Mean	SD		
Health	169	7.0059	1.94415		
Health (Dietetics)	72	6.2222	1.87062	4.178 (2, 376)	0.016
Non-health	138	6.7029	1.96840		

Table 5: Post Hoc Test between Field of Study and Eating Practices of Participants

Pair	Mean Differences (I-J)	Std. Error	Sig	95% Confidence Interval	
				Lower Bound	Upper Bound
Health, Health (Dietetics)	.78369	.27294	.013	.1273	1.4400
Health, Non-health	.30302	.22251	.522	-.2321	.8381
Health (Dietetics), Health	-.78369	.27294	.013	-1.4400	-.1273
Health (Dietetics), Non-health	-.48068	.28195	0.267	-1.1587	.1973
Non-health, Health	-.30302	.22251	0.522	-.8381	.2321
Non-health, Health (Dietetics)	.48068	.28195	0.267	-.1973	1.1587

DISCUSSION

The result obtained revealed that more than half of the participants demonstrated a good level of knowledge, suggesting that IIUM students are well-informed about key nutrition concepts, including dietary recommendations, nutrient content in food, food choices, and the relationship between diet and health.

These findings are consistent with a previous study by Yusni et al. (2023), which showed that 95.3% of participants among University Sains Malaysia (USM) students have good nutrition knowledge. In contrast, other studies have reported a much higher prevalence of poor nutrition knowledge among university students. For instance, in Kuwait, 84.1% of students had poor nutrition knowledge (Husain et al., 2021). In contrast, a UK-based study found that fewer than half of the students demonstrated an adequate understanding of nutrition (Belogianni et al., 2022).

A large proportion of participants in this study were from health-related programmes, including dietetics. Students from these fields often exhibit a higher level of nutrition knowledge due to the integration of nutrition education into their core academic syllabus. Typically, students learn about nutrition and diseases, as well as the concepts of macronutrients and micronutrients.

Moreover, the results suggested that IIUM students generally engage in moderately unhealthy dietary practices. The results are similar to those found by Ali & Jaafar (2020), who also discovered that IIUM students engaged in unhealthy dietary behaviours. This includes having fewer than three meals per day, skipping breakfast, and snacking between meals.

One of the factors that likely contributes to these dietary practices is tight academic schedules. University life often demands that students multitask between classes, assignments, examinations, and extracurricular activities, leading them to skip meals or choose meals that are easy and quick to prepare(Deliens et al., 2014). Ready-to-eat foods are often processed and have lower nutritional value.

These findings reflect a broader trend seen across university populations, where students tend to neglect and deprioritise their nutritional needs, due to challenges associated with young adulthood (Das & Evans, 2014).

The field of study has a significant impact on nutrition knowledge, with dietetics students achieving the highest scores, followed by students in other health-related fields and non-health students.

This underscores the importance of academic exposure in shaping students' knowledge. Dietetics programmes provide comprehensive nutrition education, equipping students with both theoretical and practical skills to apply evidence-based nutrition knowledge in both academic and real-world settings.

In contrast, non-health students often lack formal nutrition education, relying instead on informal sources such as social media, peers and family (Geist et al., 2024). Without access to reliable, evidence-based guidance, non-health students are more susceptible to misconceptions and fad diets, which may lead to poor dietary habits and nutrient deficiencies.

These results align with previous studies that emphasise the impact of educational background on health-related literacy. A study by Egg et al. (2020) found that increased engagement in nutrition education leads to improved nutrition knowledge and literacy, which in turn enhances eating habits.

Additionally, the dietetics students also had significantly healthier eating practices compared to students in other health-related and non-health-related fields, as reflected by their lowest mean scores in the STC questionnaires. Interestingly, the non-health students reported healthier eating practices compared to other health-related students.

The finding aligns with studies by Jahan et al. (2023) and Bayomy et al. (2024), which suggested that health and medical students often exhibit less balanced diets compared to non-health students. While exposure to health-related information might be expected to encourage healthier practices, academic stress and workload may counteract these benefits.

A study by Omar et al. (2024) reported that 35.2% of medical students in Malaysia experienced stress. Similarly, research conducted in Karachi, Pakistan, found that 75.6% of medical students attributed their high stress levels to academic demands, which is more than double the rate of students from other academic fields (Aamir, 2017). Stress arising from academic challenges, such as examination pressure and heavy workload (Othman et al., 2013), can disrupt eating behaviours, leading to emotional eating, appetite loss, or irregular meal patterns.

Poor time management further contributes to poor eating habits. Health-based students' strong dedication to studies and clinical rotations often leads them to deprioritise meal planning (Bayomy et al., 2024).

Overall, these findings highlight that nutrition knowledge is not the sole determinant of healthy eating. Psychological health, time management, financial circumstances, and environmental influences all interact to shape students' dietary behaviours. Therefore, a multidimensional approach is crucial to promoting healthier eating habits among university students.

CONCLUSION

Most students demonstrated good to excellent levels of nutrition knowledge, with the highest scores observed among those in the dietetics program. Similarly, dietetics students exhibited the healthiest eating practices compared to those in non-health-related and health-related fields.

These findings suggest that specialised nutrition education is linked to both higher knowledge and healthier eating practices.

Thus, this study highlights the importance of expanding and enhancing nutrition education beyond knowledge sharing by integrating behavioural and environmental factors in the interventions.

ACKNOWLEDGEMENT

This research is not funded by any grant.

REFERENCE

Aamir, I. S. (2017). Stress Level Comparison of Medical and Nonmedical Students: A Cross Sectional Study done at Various Professional Colleges in Karachi, Pakistan. *Acta Psychopathologica*, *03*(02). https://doi.org/10.4172/2469-6676.100080

Ali, N. K. M. F. @ M., & Jaafar, N. H. (2020). Dietary Habit and Lifestyle Practices Among Normal and Overweight/Obese IIUM Kuantan Students: A Comparative Study. *International Journal of Allied Health Sciences*, *4*(3), 1309–1320. https://doi.org/10.31436/IJAHS.V4I3.489

Bayomy, H. E., Alruwaili, S. M., Alsayer, R. I., Alanazi, N. K., Albalawi, D. A., Al Shammari, K. H., & Moussa, M. M. (2024). Eating habits of students of health colleges and non-health colleges at the Northern Border University in the Kingdom of Saudi Arabia. *PLOS ONE*, *19*(10), e0312750. https://doi.org/10.1371/JOURNAL.PONE.0312750

Belogianni, K., Ooms, A., Lykou, A., & Moir, H. J. (2022). Nutrition knowledge among university students in the UK: a cross-sectional study. *Public Health Nutrition*, *25*(10). https://doi.org/10.1017/S1368980021004754

Cena, H., & Calder, P. C. (2020). Defining a Healthy Diet: Evidence for the Role of Contemporary Dietary Patterns in Health and Disease. *Nutrients 2020, Vol. 12, Page 334*, *12*(2), 334. https://doi.org/10.3390/NU12020334

Das, B. M., & Evans, E. M. (2014). Understanding weight management perceptions in first-year college students using the health belief model. *Journal of American College Health*, *62*(7), 488–497. https://doi.org/10.1080/07448481.2014.923429,

Deliens, T., Clarys, P., De Bourdeaudhuij, I., & Deforche, B. (2014). Determinants of eating behaviour in university students: A qualitative study using focus group discussions. *BMC Public Health*, *14*(1). https://doi.org/10.1186/1471-2458-14-53,

Egg, S., Wakolbinger, M., Reisser, A., Schätzer, M., Wild, B., & Rust, P. (2020). Relationship between nutrition knowledge, education and other determinants of food intake and lifestyle habits among adolescents from urban and rural secondary schools in Tyrol, Western Austria. *Public Health Nutrition*, *23*(17), 3136. https://doi.org/10.1017/S1368980020000488

Geist, C. H., Hildebrand, D., Keirns, B. H., & Emerson, S. R. (2024). Survey of Nutrition Knowledge, Attitudes, and Preferred Informational Sources among Students at a Southwestern University in the United States: A Brief Report. *Dietetics*, *3*(2), 170–178. https://doi.org/10.3390/DIETETICS3020014

Habiba, U., & Koli, F. S. (2024). The Mediating Role of Students' Health Information Literacy Skills: Exploring the Relationship Between Web Resource Utilization and Health Information Evaluation Proficiency. *Health Expectations*, *27*(4). https://doi.org/10.1111/HEX.14176

Husain, W., Ashkanani, F., & Al Dwairji, M. A. (2021). Nutrition Knowledge among College of Basic Education Students in Kuwait: A Cross-Sectional Study. *Journal of Nutrition and Metabolism*, *2021*, 5560714. https://doi.org/10.1155/2021/5560714

Kliemann, N., Wardle, J., Johnson, F., & Croker, H. (2016). Reliability and validity of a revised version of the General Nutrition Knowledge Questionnaire. *European Journal of Clinical Nutrition*, *70*(10). https://doi.org/10.1038/ejcn.2016.87

Knowledge and Practice of Healthy Lifestyle and Dietary Habits In Medical Students: A Study From Oman. (2023). *Journal of Community Medicine and Public Health Reports*. https://doi.org/10.38207/JCMPHR/2023/SEP040503110

Mahoney, M. R., Sommer, E. C., Popescu, F., Adams, L. E., & Barkin, S. (2024). Comparing Family Health Before and After a Family-Focused Nutrition Program during the Pandemic. *Maternal and Child Health Journal*, *28*(8), 1413–1421. https://doi.org/10.1007/S10995-024-03934-2/TABLES/4

Manan, N. A., Hazwani, M., Rahimi, M., Rashidi Mohd 'asri, M., Yusuf, R. K., Faeza, S., Fuzi, A., Manan, N. A., & Author, C. (2023). Content Validity of the General Nutrition Knowledge Questionnaire for Malaysian Adult Population. In *Asian Journal of Medicine and Health Sciences* (Vol. 6). https://www.researchgate.net/publication/372395247

Moschonis, G., & Trakman, G. L. (2023). Overweight and Obesity: The Interplay of Eating Habits and Physical Activity. *Nutrients*, *15*(13), 2896. https://doi.org/10.3390/NU15132896

Omar, E., Yasin, M. M., Ismail, I. A., Shaari, S. A., Baharuddin, H., & Siran, R. (2024). Stress, Anxiety and Depression of Undergraduate Medical Students in a Public Medical Programme in Malaysia: A Risk Factor Analysis. *Education in Medicine Journal*, *16*, 37–45. https://doi.org/10.21315/EIMJ2024.16.S1.5

Ong, R. H. S., Chow, W. L., Cheong, M., Lim, G. H., Xie, W., Baggs, G., Huynh, D. T. T., Oh, H. C., How, C. H., Tan, N. C., Tey, S. L., & Chew, S. T. H. (2021). Associations between socio-demographics, nutrition knowledge, nutrition competencies and attitudes in community-dwelling healthy older adults in Singapore: findings from the SHIELD study. *Journal of Health, Population and Nutrition*, *40*(1). https://doi.org/10.1186/s41043-021-00277-4

Othman, C. N., Farooqui, M., Yusoff, M. S. B., & Adawiyah, R. (2013). Nature of Stress among Health Science Students in a Malaysian University. *Procedia - Social and Behavioral Sciences*, *105*, 249–257. https://doi.org/10.1016/J.SBSPRO.2013.11.026

Paxton, A. E., Strycker, L. A., Toobert, D. J., Ammerman, A. S., & Glasgow, R. E. (2011). Starting The Conversation: Performance of a Brief Dietary Assessment and Intervention Tool for Health Professionals. *American Journal of Preventive Medicine*, *40*(1), 67–71. https://doi.org/10.1016/J.AMEPRE.2010.10.009

Quaidoo, E. Y., Ohemeng, A., & Amankwah-Poku, M. (2018). Sources of nutrition information and level of nutrition knowledge among young adults in the Accra metropolis. *BMC Public Health*, *18*(1). https://doi.org/10.1186/s12889-018-6159-1

Serban, D. M., Ursoniu, S., Moleriu, R. D., Banu, A. M., & Serban, C. L. (2024). Mindful Eating, Nutrition Knowledge, and Weight Status among Medical Students: Implications for Health and Counseling Practices. *Nutrients 2024, Vol. 16, Page 1894*, *16*(12), 1894. https://doi.org/10.3390/NU16121894

Sukys, S., Cesnaitiene, V. J., & Ossowsky, Z. M. (2017). Is Health Education at University Associated with Students' Health Literacy? Evidence from Cross-Sectional Study Applying HLS-EU-Q. *BioMed Research International*, *2017*, 8516843. https://doi.org/10.1155/2017/8516843

Yandutkina, K., Nukeshtayeva, K., Bolatova, Z., Dauletkaliyeva, Z., Kayupova, G., Takuadina, A., Yerdessov, N., Zhamantayev, O., Zhanalina, G., Shintayeva, N., & Rogova, S. (2024). Determinants of Health Behaviors Among University Students: Insights from a Cross-Sectional Study in the Karaganda Region. *Journal of Clinical Medicine of Kazakhstan*, *21*(4), 59–65. https://doi.org/10.23950/JCMK/14966

Yusni, A., Vanoh, D., & Mohd Salleh, N. S. A. (2023). The Relationship Between Nutrition Knowledge, Attitude and Practice (NKAP), Self-Efficacy of Healthy Eating with Body Mass Index (BMI) Status Among Undergraduate Students in USM Kubang Kerian, Kelantan: A Cross-Sectional Study. *Asian Journal of Medicine and Biomedicine*, *7*(2). https://doi.org/10.37231/ajmb.2023.7.2.612

REVIEW ARTICLE

Healthy Maternal Nutrition for Child Stunting Prevention and its Promotion Through Animated Video-based Education: A Narrative Review

Amira Farhana Azizan[1], Muhamad Ashraf Rostam[1,2], Mohd Ramadan Ab Hamid[3], Muhamad Sadry Abu Seman[4], Rahiwan Nazar Romli[5], Nor Azwani Mohd Shukri[1,2,*]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM), 25200 Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, IIUM, 25200 Pahang, Malaysia
[3]Centre of Nutrition & Dietetics, Faculty of Health Sciences, Universiti Teknologi MARA, 42300 Selangor, Malaysia
[4]Kulliyyah of Information and Communication Technology, IIUM, 53100 Selangor, Malaysia
[5]Faculty of Computing, College of Computing & Applied Sciences, Universiti Malaysia Pahang Al-Sultan Abdullah, 26300 Pahang, Malaysia

Keywords:
Child stunting; exclusive breastfeeding; maternal nutrition; nutrition education; video animation
Eg: eating behaviour; picky eaters; nutritional status; children

ABSTRACT

Background: Child stunting remains a persistent public health challenge in Malaysia, despite ongoing efforts to improve maternal and child nutrition. Evidence suggests that maternal nutrition during pregnancy and exclusive breastfeeding plays a pivotal role in shaping fetal growth and early childhood development. This narrative review aimed to synthesize current evidence on maternal dietary requirements and explores the potential of animated video interventions as a strategy for nutrition education and stunting prevention. **Methods:** A literature search was conducted using PubMed, Scopus, ScienceDirect, and Google Scholar, focusing on studies published from 2000 onwards. Keywords included maternal nutrition, stunting, energy and protein intake, micronutrients, and video-based education. Inclusion criteria targeted peer-reviewed studies addressing maternal nutrition during pregnancy and the use of digital tools for health education. Thematic analysis was applied to organize findings into three domains: nutritional determinants of stunting, maternal knowledge and practices, and animated video interventions. **Results:** The findings were thematically organized into three domains: maternal nutrition requirements, maternal knowledge and practices, and the use video-based tools for health education. Studies indicate that inadequate maternal intakes of energy, protein, iron, calcium, folic acid, and vitamins A, C, and D, are associated with intrauterine growth restriction, low birth weight, and small-for-gestational-age outcomes, the key contributors to stunting. Despite national guidelines, gaps in maternal nutrition literacy and practice persist. Animated video interventions have shown potential in improving maternal understanding, engagement, and behavior change. Studies demonstrate their effectiveness in enhancing maternal knowledge and dietary practices. **Conclusion:** Maternal nutrition during pregnancy and exclusive breastfeeding is vital to child growth, with deficiencies in energy, protein, and key micronutrients linked to increased stunting risk. Despite existing guidelines, gaps in maternal knowledge persist. Animated video interventions offer a solution to improve maternal nutrition literacy and support behavior change, with potential to enhance stunting prevention outcomes.

INTRODUCTION

Child stunting, characterised by impaired linear growth due to chronic undernutrition, recurrent infections, and insufficient psychosocial stimulation (World Health Organisation (WHO), 2015) remains a pressing public health concern in Malaysia. Despite global progress in reducing stunting ratesfrom 40% in 1990 to 25% in 2013 (de Onis et al., 2016),Malaysia has experienced a troubling upward trend. Between 2012 and 2022, the prevalence of stunting among children under five rose from 17.6% to 21.9% (United Nations Children's Fund (UNICEF)/WHO/World Bank, 2023), with particularly high rates in the East Coast region, including Kelantan (34%), Terengganu (26.1%), and Pahang (25.7%) (Haron et al., 2023). This divergence from global patterns signals a need for targeted, context-specific interventions.

The first 1,000 days of life, which include the timeframe from conception up until the age of two years, constitute a crucial period for growth and development. Maternal nutrition during pregnancy plays a pivotal role in shaping foetal outcomes and preventing early childhood stunting. Studies have shown that maternal undernutrition contributes to approximately 20% of stunting cases (Özaltin et al., 2010), with additional risk factors including micronutrient deficiencies, poor dietary diversity, and limited nutrition literacy (Black et al., 2013; Morrison & Regnault, 2016). Inadequate maternal nutrition not only compromises birth outcomes but also affects long-term cognitive development and economic productivity (Mustakim et al., 2022; McGovern et al., 2017).

Despite the availability of national guidelines and supplementation programs, gaps persist in maternal knowledge and practice. Research indicates that many mothers lack awareness of trimester-specific nutrient needs (Renuka et al., 2020; Nana & Zema, 2018). Traditional education toolssuch as pamphlets and face-to-face counselling,often fail to engage diverse populations effectively, especially in rural or underserved areas. Meanwhile, digital health strategies have gained traction

as scalable, engaging platforms for nutrition education. Animated videos, grounded in Cognitive Load Theory, offer a promising medium for simplifying complex information and enhancing retention (Ab Hamid, 2022; Brame, 2016). Studies have demonstrated that video-based interventions can improve maternal knowledge, attitudes, and practices related to maternal nutrition and complementary feeding (Ramadhani et al., 2023; Schneider et al., 2022; Rodrigues et al., 2017).

This narrative review synthesizes current evidence on maternal nutrition during pregnancy and explores the potential of animated video interventions as a novel strategy for stunting prevention. By integrating nutritional science with digital innovation, this review aims to inform future program design and policy development in maternal and child health.

MATERIALS AND METHODS

Literature Search Strategy

A systematic search was conducted across multiple databases including PubMed, Scopus, ScienceDirect, and Google Scholar. Among the main keywords used included 'maternal nutrition', 'total energy', 'macronutrients', 'micronutrient', 'pregnancy', 'stunting', 'nutrition education', 'video', and 'animation'. Several additional search terms recognized as risk factors for stunting, such as 'low birth weight', 'small-for-gestational age', and 'intrauterine growth retardation' were also used. Boolean operators and truncation were applied to refine the results. The reference lists of selected articles were also scrutinized for any other relevant studies.

Inclusion and Exclusion Criteria

The findings were further refined based on the pre-determined inclusion and exclusion criteria (Table 1). A restriction of date was put on the search where only the year 2000 onwards were included because this timeframe was determined in consideration of the fact that this was the period in which research on information and communication technology adoption in the digital health environment began to be conducted globally. Articles that were not written in Malay or English were excluded from this study. As this review focused primarily on maternal nutrition during pregnancy, articles about non-pregnant mothers were automatically excluded during the initial screening. The search criteria were broadened to encompass the relationship between maternal nutrition and the prevention of stunting and the effectiveness of video animation in delivering the information to expand the range of articles available for review.

Table 1: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
1. Studies published from year 2000 onwards.	1. Studies published before year 2000
2. Peer-reviewed journal articles, reports, and guidelines published in English or Malay.	2. Non-English or non-Malay publications.
3. Studies focusing on maternal nutrition during pregnancy.	3. Studies not directly related to maternal nutrition or stunting.
4. Studies evaluating digital or video-based nutrition education tools.	4. Studies focused solely on infant feeding without maternal context.

Data Extraction and Thematic Analysis

Key information from the selected studies was organized and synthesized according to common themes. Deductive approach was used to determine the appropriate themes that were relevant to the research questions and aims of the review based on existing findings. The themes were focused on maternal nutrition needs, maternal behaviors, and the use of digital tools like animated videos for nutrition education.

Ethical Considerations

As this review utilized secondary data from published sources, no ethical approval was required. However, all sources were appropriately cited and referenced in accordance with academic integrity standards.

RESULTS AND DISCUSSION

In total, approximately 18 peer-reviewed articles were identified and examined for this review. These publications included both primary research and review papers addressing maternal nutrition requirements during pregnancy, maternal knowledge and practices, and the use of video-based tools for health education. Rather than presenting each article individually, the evidence was synthesised thematically to highlight key findings and trends across the literature (Table 2). Key studies most relevant to Malaysian and regional contexts were emphasised to provide depth and applicability to the discussion. Citations within the discussion are used to illustrate key insights and do not represent an exhaustive listing of all reviewed sources.

Table 2. Key insights from the review according to themes

Theme	Representative Articles	Key Insight
Maternal nutrition during pregnancy	González-Fernández et al. (2024) Mohamed et al. (2022) Eshak et al. (2022) Yang et al. (2022) Souza et al. (2021) Miliku et al. (2016)	Inadequate intakes of energy, protein, and several micronutrients have been associated with key risk factors for child stunting.
Maternal nutrition knowledge and practices	Nguyen et al. (2017) Nana and Zema (2018) Renuka et al. (2020)	Despite availability of guidelines, maternal nutrition literacy is still inadequate and does not consistently translate into practices, highlighting the need for more effective education strategies.
Videos as health education tools	Ab Hamid et al. (2022) Collishaw et al. (2023) Schneider et al. (2022) Ramadhani et al. (2023)	Videos designed by considering factors that capture attention and enhance information retention can provide an effective approach to delivering maternal nutrition education.

Maternal Nutrition During Pregnancy

Children's linear growth is particularly vulnerable to environmental stressors and malnutrition during the first 1,000 days of life, from conception to their second birthdays. During the first 500 days, from conception to about six months of age, the child is entirely dependent on the mother for its nutrition, either via the placenta during pregnancy or via breastmilk during the initial six-month exclusive breastfeeding period (Aguayo & Menon, 2016). Therefore, maternal nutrition is a cornerstone of fetal development and early childhood growth. Evidence from studies have demonstrated that an adequate intake of energy, protein, and selected micronutrients during pregnancy and exclusive breastfeeding to prevent intrauterine growth restriction and postnatal stunting (González-Fernández et al., 2024).

Total Energy and Protein

Poor maternal nutritional status such as low body mass index as well as inadequate energy and nutrients intakes before and during pregnancy can lead to intrauterine growth restriction (IUGR) (Mohamed et al., 2022; Clark, 2018), resulting in small-for-gestational age (SGA) birth weight and reduced birth length, which are among the key risk factors for stunting in Malaysia (Makbul et al., 2025). On the other hand, maternal total energy intake was positively associated with the offspring's birth weight in a national cohort study involving 78,793 Japanese mothers (Eshak at al., 2020). The total energy intake for pregnant women is slightly higher than that of their non-pregnant counterparts to support optimal gestational weight gain (Most et al., 2019). According to the Nutrition Division (Ministry of Health (MOH), 2022), pregnant mothers need 2,000–2,470 kcal per day for healthy fetal growth, while lactating mothers need 2,500 kcal per day. On the other hand, according to the Recommended Nutrient Intakes for Malaysia (RNI) (National Coordinating Committee on Food and Nutrition (NCCPN), 2017) and Perinatal Care Manual (MOH, 2022), each trimester of pregnancy requires different energy intake to maintain the fetus's healthy growth. Both guidelines advised adding 80 kcal, 280 kcal, and 470 kcal during the first, second, and third trimesters. After delivery, the RNI suggests taking an additional 450 kcal, while the Perinatal Care Manual advises adding 500 kcal from a daily intake of 1,600 kcal to 1,800 kcal to meet the increased requirements of lactation.

Adequate maternal protein intake during pregnancy is associated with healthy birth weight and length. Deficiencies, particularly in animal and dairy protein, have been linked to IUGR, low birth weight (LBW) and SGA (Clark, 2018). On the other hand, higher intake of dietary protein, in particular animal protein and dairy protein, is associated with higher birth weight and lower risks of LBW, SGA, and IUGR (Yang et al., 2022). A cross-sectional study on 7,750 pregnant women in China demonstrated that every 3% increase in calories from total protein, animal protein, and dairy protein was linked to increases in birth weight of 19.4 g (95% CI 6.0–32.9), 20.6 g (4.8–36.5), and 18.2 g (4.7–31.7), respectively. Adequate protein intake is important as it involves structural and functional biological roles. A systematic review by Mohd Shukri et al. (2022) found that protein intake between 0.8–1.1 g/kg body weight and consumption of at least five food groups daily were linked to improved birth weight and length. Pregnant mothers need 62.5 g of protein throughout pregnancy. On the other hand, *Bahagian Pemakanan dan Kesihatan Keluarga* (BPKK) suggests adding 0.5 g, 8 g, and 25 g of protein during the first, second, and third trimesters, respectively, to achieve the daily protein requirement (MOH, 2022). After delivery and exclusive breastfeeding, an additional 19 g of protein is considered sufficient.

Micronutrients

Micronutrients are chemical substances that, when consumed in minimal amounts, enable the regulation of metabolic and biochemical functions of the body. The requirement for micronutrients, which collectively comprise vitamins and minerals, also increases during pregnancy. This is because they are crucial to the growth and development of the unborn child in the womb as well as the health of expectant mothers. A few vitamins and minerals deficiencies in mothers have been associated with the prevalence of low birth weight and SGA in their babies such as vitamin D, iron and calcium (González-Fernández et al., 2024).

Iron is a vital mineral that is integral to haemoglobin, thereby facilitating oxygen transfer. Iron requirements during pregnancy increase dramatically as the mother's blood volume expands and the fetus grows and develops (Georgieff, 2020). In fact, the most prevalent nutritional deficiency among expectant mothers is iron deficiency anemia. During pregnancy, 30−60 mg of iron is required to avoid anemia (NCCPN, 2017). Examples of iron-rich foods are liver, lean meats, egg yolks, legumes (such as chickpeas and peas), and dried fruits (raisins and dates), as well as iron-fortified foods (iron-fortified bread and breakfast cereals). Iron deficiency is associated with adverse birth outcomes including low birth weight, intrauterine growth restriction and premature delivery (Georgieff, 2020).

Maternal Vitamin C or ascorbic acid intake during pregnancy has been associated with increased birth weight at term in a cross-sectional study involving 626 mother-infant pairs (Souza et al., 2021). Vitamin C supports antioxidant defenses against oxidative stress during pregnancy. These defenses protect maternal and fetal tissues from damage, and their imbalance can impair outcomes such as fetal growth. It also serves as a co-factor in collagen synthesis, supporting cartilage and bone development in the newborn (Jang et al., 2018). The requirement for vitamin C during pregnancy is 80 mg (NCCPN, 2017). Food sources that are easily available and contain high levels of vitamin C, such as guava, lemon, and papaya, should be taken by pregnant or lactating mothers.

Vitamin D is a fat-soluble vitamin that plays an important role in maintaining musculoskeletal health. During pregnancy, the foetus is totally dependent on maternal supply of vitamin D. A review on 43 review articles, and 31 original articles found that a pregnancy-related 25(OH)D deficit would raise the risk of preterm birth, LBW, and SGA (Tsenkova-Toncheva et al., 2024). Miliku and colleagues (2016) conducted a prospective cohort study in Rotterdam, Netherlands, and discovered that offspring of mothers with lower 25(OH)D levels had increased foetal growth restriction in the third trimester. This restriction resulted in a significantly smaller head circumference, shorter body length, and a lower birth weight compared to mothers in the highest quartile. Pregnant mothers require a daily dose of 15 µg of vitamin D (NCCPN 2017). Food sources rich in vitamin D are sardines, eggs, meat, and mackerel. Exposure to the sun also helps mothers get enough vitamin D every day.

A systematic review on 30 articles indicates that calcium is among the nutrients that affect birth weight (Mousa et al., 2019). As for calcium requirements, pregnant mothers need 1,000 mg daily to form the baby's bones and teeth, especially in the third trimester (NCCPN, 2017; MOH, 2022). It also helps prevent muscle spasms in pregnant women. This requirement is 200 mg higher than for women who are not pregnant (NCCPN, 2017). Examples of foods high in calcium are milk and dairy products, anchovies, sardines, beans, and green leafy vegetables. Mothers should drink two glasses of milk daily to complete the calcium requirement.

Maternal Nutrition Knowledge and Practices

Despite the availability of national guidelines, gaps persist in maternal nutrition literacy. Studies show that while many mothers possess basic dietary knowledge, fewer translate it into practice. Based on a study by Nguyen et al. (2017), the intake of iron and folic acid supplementation showed a significant gap in the knowledge and practice of pregnant mothers, where at least 20% of women in the childbearing period are underweight. This study proves that there are still a high number of people who lack the self-awareness to start healthy practices during pregnancy and breastfeeding. Therefore, the delivery of maternal interventions should be improved to reduce the high burden of maternal and child undernutrition and mortality (Bhutta et al., 2013).

In addition, Nana and Zema (2018) found that only 39.3% of pregnant women practiced good dietary habits despite over half demonstrating adequate knowledge. On the other hand, a study was carried out by Renuka and colleagues (2020) on the effectiveness of educational interventions on breastfeeding among pregnant women. They gave a booklet to pregnant mothers and later followed them up regarding knowledge, attitude, and practices during the first week of delivery, the third month, and the sixth month of post-delivery. The results showed that only 2.8% of the participants knew that breastfeeding should begin within one hour of birth, and more than half did not know the right time to start breastfeeding after delivery. In addition, just 22.8% of the women understood

the importance of exclusive breastfeeding for six months. These findings underscore the need for more engaging, accessible education strategies to bridge the knowledge–practice gap.

Videos as Health Education Tools

Digital transformation for delivering health information must be done to achieve effectiveness in health education (WHO, 2021). These worldwide plans aim to maintain digital health technologies aligned with the national health sector and health strategies. It is anticipated that by 2025, a digital health ecosystem will be established in the country, encompassing a variety of digital information technology infrastructure primarily utilized by healthcare providers, health service providers, and patients (WHO, 2021). Public health authorities, universities, and research institutions are also expected to use this infrastructure. Videos and images can attract viewer interest, enhance learning, and reach more people worldwide without boundaries or limitations. In contrast, traditional nutrition education tools (face-to-face counseling, printed booklets, and pamphlets) have limited outreach and can only be targeted at certain places.

Digital health tools, particularly animated videos, offer a promising solution to enhance maternal nutrition education. Nowadays, video education is more practical than other hardcopy tools such as pamphlets, flyers, or newspapers. This is because animated videos can simplify complex information, reduce extraneous cognitive burden, and improve retention. According to the Cognitive Load Theory, there are some significant extra values in educational videos that can attract viewers' interest. These include signaling, segmenting, weeding, and modality matching (Brame, 2016).

Signaling, or cueing, effectively uses on-screen text or signs to draw attention to important information and enhance information retention. This can be achieved through use of keywords, a change in color or contrast, or a symbol like an arrow that guides the viewer's gaze to a specific area of the screen (Brame, 2016). Emphasizing critical information helps the learner's explicit engagement and reduces the extraneous load. Research by Clark & Mayer (2016) has indicated that this approach enhances students' capacity to retain and apply new knowledge acquired from animations.

To create an impactful video, segmenting information into smaller chunks is important. This segmentation allows learners to engage with new information at their own pace and retain it more effectively. Additionally, it is beneficial to eliminate distracting elements such as music, complex backgrounds, or unnecessary features that might divert the learner's attention. This video animation gives them better control over the flow of information and helps them focus on what is important. In fact, the study conducted by Ab Hamid and team (2022) revealed that the developed videos on dietary management are highly understandable, actionable, and suitable for use as educational videos.

In certain countries, nurses have used educational videos as an effective education and health promotion strategy due to their versatility and applicability (Rodrigues et al., 2017). In brief, the combination of educational videos and interventions by healthcare professionals is thought to enhance public understanding of the issues, leading to the early identification and resolution of health problems (Rodrigues et al., 2017). As determined by Collishaw et al. (2023), students exposed to video treatment improved their nutrition knowledge scores compared to those who received traditional classroom instruction. The above studies proved that education can be delivered more conveniently and practically through video to reach more people effectively.

Referring to Schneider et al. (2022), 12 community health workers (CHW) were given educational videos on nutrition based on the WHO's recommendation to support their work during their visits with the families. Three months later, they conducted interviews individually with eight CHW and 16 mothers to investigate the effectiveness of the education videos in delivering information. In fact, the study revealed that learning from nutrition education videos was easier compared with traditional education and showed improvement in their daily practices.

In relation to the study by Ramadhani et al. (2023), education using animated videos and leaflets has a significant effect on increasing the knowledge of mothers who had babies aged 4–5 months to prepare complementary feeding. The researchers conducted pre- and post-tests to gauge the mother's understanding of how to prepare complementary feeding for stunting prevention. The results showed significant improvements in the knowledge and attitude of mothers after an animated video intervention on complementary feeding.

In Malaysia, the digital health ecosystem is expanding, with national strategies aiming to integrate digital tools into maternal and child health services by 2025 (WHO, 2021). Leveraging animated videos within this framework could enhance outreach, especially in underserved regions with limited access to face-to-face counseling.

As a narrative review, this study is limited by the absence of a systematic search strategy and formal quality

appraisal, which may introduce selection bias and restrict the reproducibility of findings. Additionally, the lack of quantitative synthesis limits the ability to draw definitive conclusions on the magnitude of associations reported. However, despite these limitations, the review provides a valuable overview to inform future research and intervention design in maternal nutrition and stunting prevention by leveraging digital platforms for health education.

CONCLUSION

Maternal nutrition during pregnancy and exclusive breastfeeding is essential for optimal fetal growth and early childhood development. Deficiencies in energy, protein, and selected micronutrients such as iron, calcium, folic acid, and vitamins A and D, are associated with key risk factors for child stunting. Despite existing guidelines, gaps in maternal knowledge and practice persist. Animated video interventions offer a promising approach to enhance nutrition literacy and support behavior change. Integrating such tools into maternal health programs may strengthen child stunting prevention efforts and improve long-term outcomes.

ACKNOWLEDGEMENT

The authors would like to express their appreciation to every researcher whose work has been cited in this review. This research was supported by the IIUM-UMP-UITM Sustainable Research Collaboration Grant 2020 (SRCG20-005-0005).

REFERENCES

Aguayo, V. M., & Menon, P. (2016). Stop stunting: improving child feeding, women's nutrition and household sanitation in South Asia. Maternal & child nutrition, 12 Suppl 1(Suppl 1), 3–11. https://doi.org/10.1111/mcn.12283

Ab Hamid, M.R., Mohd Yusof, N.D., Buhari, S.S. (2022). Understandability, actionability and suitability of educational videos on dietary management for hypertension. Health Educ J 81(2):238–247. https://doi.org/10.1177/00178969211072275

Bhutta, Z. A., Das, J. K., Rizvi, A., Gaffey, M. F., Walker, N., Horton, S., Webb, P., Lartey, A., Black, R. E., & Lancet Nutrition Interventions Review Group, the Maternal and Child Nutrition Study Group (2013). Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost?. Lancet (London, England), 382(9890), 452–477. https://doi.org/10.1016/S0140-6736(13)60996-4

Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., de Onis, M., Ezzati, M., Grantham-McGregor, S., Katz, J., Martorell, R., Uauy, R., & Maternal and Child Nutrition Study Group (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. Lancet (London, England), 382(9890), 427–451. https://doi.org/10.1016/S0140-6736(13)60937-X

Brame, C. J. (2016). Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content. CBE life sciences education, 15(4), es6. https://doi.org/10.1187/cbe.16-03-0125

Clark, D. C. (2018). Association of Dairy Protein Intake During Pregnancy with Birth Weight. Food and nutrition bulletin, 39(2_suppl), S54–S59. https://doi.org/10.1177/0379572118775824

Clark, R. C., & Mayer, R. E. (2016). E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning. Wiley

Collishaw, A., Snider, A., McNamara, P., Nuvaga, S., & Bilame, A. F. (2024). Impact of a video-based nutrition education program on the nutrition knowledge of students and parents: evidence from the North Region of Cameroon. Global health promotion, 31(3), 60–69. https://doi.org/10.1177/17579759231206795

de Onis, M., & Branca, F. (2016). Childhood stunting: a global perspective. Maternal & child nutrition, 12 Suppl 1(Suppl 1), 12–26. https://doi.org/10.1111/mcn.12231

Eshak, E. S., Okada, C., Baba, S., Kimura, T., Ikehara, S., Sato, T., Shirai, K., Iso, H., & Japan Environment and Children's Study Group (2020). Maternal total energy, macronutrient and vitamin intakes during pregnancy associated with the offspring's birth size in the Japan Environment and Children's Study. The British journal of nutrition, 124(6), 558–566. https://doi.org/10.1017/S0007114520001397

Georgieff M. K. (2020). Iron deficiency in pregnancy. American journal of obstetrics and gynecology, 223(4), 516–524. https://doi.org/10.1016/j.ajog.2020.03.006

González-Fernández, D., Muralidharan, O., Neves, P. A., & Bhutta, Z. A. (2024). Associations of Maternal Nutritional Status and Supplementation with Fetal, Newborn, and Infant Outcomes in Low-Income and Middle-Income Settings: An Overview of Reviews.

Nutrients, 16(21), 3725. https://doi.org/10.3390/nu16213725

Haron, M. Z., Rohana, A. J., Hamid, N. A. A., Omar, M. A., & Abdullah, N. H. (2023). Stunting and Its Associated Factors among Children Below 5 Years Old on the East Coast of Peninsular Malaysia: Evidence from the National Health and Morbidity Survey. The Malaysian journal of medical sciences : MJMS, 30(5), 155–168. https://doi.org/10.21315/mjms2023.30.5.13

Jang, W., Kim, H., Lee, B. E., & Chang, N. (2018). Maternal fruit and vegetable or vitamin C consumption during pregnancy is associated with fetal growth and infant growth up to 6 months: results from the Korean Mothers and Children's Environmental Health (MOCEH) cohort study. Nutrition journal, 17(1), 105. https://doi.org/10.1186/s12937-018-0410-6

Makbul, I. A. A., Yeo, G. S., Sharif, R., Lim, S. M., Mediani, A., Geurts, J., Poh, B. K., & On Behalf Of The Seanuts Ii Malaysia Study Group (2025). Determinants of Stunting Among Children Aged 0.5 to 12 Years in Peninsular Malaysia: Findings from the SEANUTS II Study. *Nutrients*, *17*(14), 2348. https://doi.org/10.3390/nu17142348

McGovern, M. E., Krishna, A., Aguayo, V. M., & Subramanian, S. V. (2017). A review of the evidence linking child stunting to economic outcomes. International journal of epidemiology, 46(4), 1171–1191. https://doi.org/10.1093/ije/dyx017

Miliku, K., Vinkhuyzen, A., Blanken, L. M., McGrath, J. J., Eyles, D. W., Burne, T. H., Hofman, A., Tiemeier, H., Steegers, E. A., Gaillard, R., & Jaddoe, V. W. (2016). Maternal vitamin D concentrations during pregnancy, fetal growth patterns, and risks of adverse birth outcomes. The American journal of clinical nutrition, 103(6), 1514–1522. https://doi.org/10.3945/ajcn.115.123752

Ministry of Health Malaysia. (2022). Perinatal Care Manual 4th Edition. Putrajaya (KL): MoH.

Mohamed, H. J. J., Loy, S. L., Mitra, A. K., Kaur, S., Teoh, A. N., Rahman, S. H. A., & Amarra, M. S. (2022). Maternal diet, nutritional status and infant birth weight in Malaysia: a scoping review. BMC pregnancy and childbirth, 22(1), 294. https://doi.org/10.1186/s12884-022-04616-z

Mohd Shukri, N. A., Mohamed Nor, N., Mustafa, N., Azizan, A. F., Rostam, M. A., Abu Seman, M. S., Ab Hamid, M. R., & Romli, R. N. (2023). A Systematic Review of Maternal Dietary Intake and its Association with Childhood Stunting. IIUM Medical Journal Malaysia, 22(1). https://doi.org/10.31436/imjm.v22i1.2141

Most, J., Dervis, S., Haman, F., Adamo, K. B., & Redman, L. M. (2019). Energy Intake Requirements in Pregnancy. Nutrients, 11(8), 1812. https://doi.org/10.3390/nu11081812

Mousa, A., Naqash, A., & Lim, S. (2019). Macronutrient and Micronutrient Intake during Pregnancy: An Overview of Recent Evidence. Nutrients, 11(2), 443. https://doi.org/10.3390/nu11020443

Morrison, J. L., & Regnault, T. R. (2016). Nutrition in Pregnancy: Optimising Maternal Diet and Fetal Adaptations to Altered Nutrient Supply. Nutrients, 8(6), 342. https://doi.org/10.3390/nu8060342

Mustakim, M. R. D., Irwanto, Irawan, R., Irmawati, M., & Setyoboedi, B. (2022). Impact of Stunting on Development of Children between 1-3 Years of Age. Ethiopian journal of health sciences, 32(3), 569–578. https://doi.org/10.4314/ejhs.v32i3.13

Nana, A., & Zema, T. (2018). Dietary practices and associated factors during pregnancy in northwestern Ethiopia. BMC pregnancy and childbirth, 18(1), 183. https://doi.org/10.1186/s12884-018-1822-1

National Coordinating Committee on Food and Nutrition (NCCFN). (2017). Recommended Nutrient Intakes (RNI). Putrajaya (KL): Ministry of Health Malaysia.

Nguyen, P. H., Sanghvi, T., Kim, S. S., Tran, L. M., Afsana, K., Mahmud, Z., Aktar, B., & Menon, P. (2017). Factors influencing maternal nutrition practices in a large scale maternal, newborn and child health program in Bangladesh. PloS one, 12(7), e0179873. https://doi.org/10.1371/journal.pone.0179873

Özaltin, E., Hill, K., & Subramanian, S. V. (2010). Association of maternal stature with offspring mortality, underweight, and stunting in low- to middle-income countries. JAMA, 303(15), 1507–1516. https://doi.org/10.1001/jama.2010.450

Ramadhani, D.W., Setianingsih, D., Wahyuni, N. (2023). The effect of early education using animation video and leaflets on preparation of complementary feedings as stunting prevention. KnE Medicine 31(1):64−75. https://doi.org/10.18502/kme.v3i1.12699

Renuka, M., Shabadi, N., Kulkarni, P., Kumar, D.S., Anup, G., Murthy, M.N. (2020). Effectiveness of educational

intervention on breastfeeding among primi pregnant women - A longitudinal study. Clin Epidemiol Glob Health 8(4):1306-1311. https://doi.org/10.1016/j.cegh.2020.05.002

Rodrigues, J.C., Rebouças, C.B., Castro, R.C., Oliveira, P.M., Almeida, P.C., Pagliuca, L.M. (2017). Development of an educational video for the promotion of eye health in school children. Texto Contexto Enferm 26(2):e06760015. https://doi.org/10.1590/0104-07072017006760015

Schneider, L., Ollila, S., & Mutanen, M. (2022). The usefulness of nutrition and health videos displayed on mobile phones in rural Uganda: Experiences of community health workers and mothers. Maternal & child nutrition, 18(2), e13322. https://doi.org/10.1111/mcn.13322

Souza, R. C. V. E., Miranda, C., & Dos Santos, L. C. (2021). Maternal vitamin B3 and C intake in pregnancy influence birth weight at term. Nutrition (Burbank, Los Angeles County, Calif.), 91-92, 111444. https://doi.org/10.1016/j.nut.2021.111444

Tsenkova-Toncheva, L., Hristova-Atanasova, E., Iskrov, G., & Stefanov, R. (2024). Prenatal Vitamin D Deficiency and Maternal and Fetal Health Outcomes. Cureus, 16(9), e69508. https://doi.org/10.7759/cureus.69508

UNICEF/WHO/International Bank for Reconstruction and Development/The World Bank (2023). United Nations Children's Fund, World Health Organization, International Bank for Reconstruction and Development/The World Bank. Levels and trends in child malnutrition. New York (USA): UNICEF and WHO.

World Health Organization (2015). Stunting in a nutshell. https://www.who.int/news/item/19-11-2015-stunting-in-a-nutshell [Accessed 21st November 2022].

World Health Organization (2021) Global strategy on digital Health 2020–2025. Geneva (CH): World Health Organization.

Yang, J., Chang, Q., Tian, X., Zhang, B., Zeng, L., Yan, H., Dang, S., & Li, Y. H. (2022). Dietary protein intake during pregnancy and birth weight among Chinese pregnant women with low intake of protein. Nutrition & metabolism, 19(1), 43. https://doi.org/10.1186/s12986-022-00678-0

ORIGINAL ARTICLE

Sensory Evaluation of Bread Formulated with Jackfruit (*Artocarpus heterophyllus*) Flour at Varying Substitution Levels

Ibrahim Abu Bakar[1], Nurulhuda Mahasan[2,*]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
Jackfruit Flour; Functional Bread; Gluten-Reduced Bread; Sensory Evaluation; Consumer Acceptability

ABSTRACT

Background: The increasing demand for low-gluten bakery products has led to growing interest in the use of jackfruit (*Artocarpus heterophyllus*) flour as a functional ingredient in bread formulation. This study investigates the sensory properties of bread produced with varying levels of jackfruit flour substitution (0%, 20%, and 40%) to evaluate its suitability as a partial wheat flour replacement. **Methods:** Jackfruit flour was produced through cleaning, slicing, oven-drying, grinding, and sieving of unripe jackfruit, followed by incorporation into bread dough using standardized ingredient proportions and baking conditions. Sensory evaluation was conducted with 30 untrained panelists who assessed appearance, aroma, texture, taste, and overall acceptability using a 9-point hedonic scale. Data were analyzed using descriptive statistics and one-way ANOVA. **Results:** Results showed that partial substitution with 20% jackfruit flour produced bread with acceptable sensory scores comparable to the control, while 40% substitution resulted in significantly lower scores for appearance, texture, and overall acceptability ($p < 0.05$). **Conclusion:** The findings suggest that jackfruit flour can be incorporated up to 20% without compromising sensory quality, offering a promising approach for developing low-gluten, fibre-enriched bread. Further research incorporating instrumental analysis is recommended to complement sensory evaluation.

INTRODUCTION

Bread is one of the most widely consumed staple foods globally and serves as a major source of carbohydrates, protein, and dietary energy. Traditional bread relies primarily on wheat flour, which contains gluten; a structural protein essential for dough elasticity and gas retention. However, the increasing prevalence of gluten-related disorders and consumer demand for healthier alternatives have driven research on gluten-reduced and functional bakery products (Lebwohl et al., 2023; Melini & Melini, 2021). Recent dietary trends also highlight the need for low-glycaemic bakery formulations to support individuals managing diabetes and metabolic syndrome (Ferrari et al., 2022).

Jackfruit (*Artocarpus heterophyllus*) flour has gained scientific interest as a functional ingredient due to its favourable nutritional properties. Made from unripe jackfruit, the flour is rich in dietary fibre, resistant starch, and bioactive compounds, while naturally free of gluten (Kumar et al., 2023; Shilpa & Lakshmi, 2021). Studies have demonstrated its potential for improving fibre content, lowering glycaemic response, and enhancing mineral availability when incorporated into bakery and snack products (Reddy et al., 2022; Shelar et al., 2023). As an underutilised tropical crop, jackfruit also offers sustainability advantages by reducing food waste and promoting plant-based food diversification.

Despite these benefits, substituting wheat flour with jackfruit flour can influence critical sensory attributes such as texture, appearance, flavour, and aroma—factors that strongly determine consumer acceptance in bakery products (Costa et al., 2023). Previous research reports that gluten dilution and the high fibre content of jackfruit flour may affect dough handling, crumb structure, and flavour profile. However, most existing studies emphasize nutritional and functional properties rather than comprehensive sensory evaluation of jackfruit-based bread.

Therefore, this study aims to evaluate the sensory acceptability of bread formulated with jackfruit flour at substitution levels of 0%, 20%, and 40%, focusing on appearance, aroma, texture, taste, and overall acceptability using a hedonic scale. By identifying the substitution threshold that maintains desirable sensory qualities, this study provides insights for the formulation of nutritionally enhanced, gluten-reduced bread using jackfruit flour as a sustainable functional ingredient.

* *Corresponding author.*
E-mail address: iab@iium.edu.my

MATERIALS AND METHODS

Research Design

A laboratory-based experimental design was employed to examine the sensory properties of bread made with different ratios of jackfruit flour and wheat flour. All procedures—including flour preparation, ingredient blending, dough development, and baking—were standardised to ensure that differences in sensory evaluation could be attributed solely to substitution levels.

Raw Materials Preparation

Fresh unripe jackfruit, commercial wheat flour, yeast, sugar, vegetable oil, and salt were sourced from local markets. All ingredients were inspected for quality and stored under appropriate conditions to prevent contamination or moisture absorption before use.

Jackfruit Flour Production

Unripe jackfruit was cleaned, peeled, and chopped into small pieces (Figure 1).



Figure 1: Chopped jackfruit pieces prepared for drying

The pieces were arranged in a single layer and oven-dried at 60°C for 12 hours to reduce moisture content. Dried slices were ground using an electric grinder, followed by sieving through a <250 µm mesh to obtain fine, uniform flour appropriate for bakery applications (Figure 2). The flour was stored in airtight containers until used.



Figure 2: Grinding of dried jackfruit slices into flour

Preparation of Ingredient Blends

Three flour formulations were prepared:

a) Control (0%): 100% wheat flour (Sample 1)
b) 20% substitution: 80% wheat flour + 20% jackfruit flour (Sample 2)
c) 40% substitution: 60% wheat flour + 40% jackfruit flour (Sample 3)

All formulations were measured by weight using digital scales to ensure accuracy and consistency.

Bread Production

Bread dough was prepared using the following standard recipe for all samples, which were (i) 3 cups flour blend, (ii) 1 cup warm water, (iii) 2 tablespoons white sugar, (iv) 1.5 teaspoons instant yeast, and (v) 1 teaspoon salt. Ingredients were mixed until homogeneous, then proofed under identical temperature and humidity conditions. Dough samples were baked in an automatic bread maker at 180°C for 25 minutes to ensure uniform development of crust and crumb across all formulations. After baking, bread samples were cooled to room temperature and stored individually in coded resealable bags to prevent bias during sensory evaluation (Figure 3).



Figure 3: Cross-section of jackfruit bread formulated with 20% jackfruit flour substitution, showing crumb texture and colour

Experimental Variables

The experimental design incorporated both independent and dependent variables. The independent variable was the level of jackfruit flour substitution, which consisted of three formulation levels: 0%, 20%, and 40%. The dependent variables were the sensory attributes evaluated by the panelists, namely appearance, aroma, texture, and taste, along with the overall acceptability score. All sensory responses were obtained through a structured evaluation method to ensure consistency, reliability, and comparability of the data across all bread formulations.

Data Collection

Sensory evaluation was conducted with 30 untrained panelists from IIUM Kuantan (aged 20–30 years). Inclusion criteria included non-smokers, no known food allergies, and no reported taste or olfactory impairments. Bread samples were coded using random three-digit numbers and presented in randomised order to minimise psychological and positional bias.

Panelists assessed appearance, aroma, texture, taste, and overall acceptability using a 9-point hedonic scale ranging from "1 = Dislike extremely" to "9 = Like extremely." All evaluations were performed under controlled laboratory lighting and room temperature.

Statistical Analysis

All data were analysed using SPSS Version 26. Means ± standard deviation (SD) were calculated for each sensory attribute. Differences between bread formulations were tested using one-way ANOVA followed by Bonferroni post hoc tests where applicable. Statistical significance was set at $p < 0.05$. Mean sensory scores were compared to determine the most acceptable jackfruit flour substitution level. Emphasis was placed on identifying whether 20% or 40% substitution maintained acceptable sensory quality relative to the control.

RESULTS AND DISCUSSION

Statistical Analysis of Sensory Evaluation

The sensory evaluation results for appearance, aroma, texture, taste, and overall acceptability were analysed using one-way ANOVA. Significant differences were detected among the three bread formulations ($p < 0.05$), demonstrating that the level of jackfruit flour substitution influenced the sensory attributes assessed by the panel. The mean scores for each sensory parameter are presented in Table 1.

Table 1: The Mean Scores of Sensory Characteristics Between Samples

Sensory Characteristics	N	S1 (Mean ± SD)	S2 (Mean ± SD)	S3 (Mean ± SD)	p-value	Significance Level
Appearance	30	7.57 ± 1.46	6.57 ± 1.61	6.07 ± 1.79	p = 0.00	S
Aroma	30	7.50 ± 1.38	6.67 ± 1.66	5.97 ± 1.63	p = 0.00	S
Texture	30	7.63 ± 1.37	5.43 ± 1.87	4.83 ± 1.91	p = 0.00	S
Taste	30	7.70 ± 1.44	4.90 ± 1.90	4.47 ± 1.91	p = 0.00	S
Overall Acceptance	30	7.70 ± 1.37	5.60 ± 1.71	4.87 ± 1.78	p = 0.00	S

Note: S = Significant ($p < 0.05$)

Appearance of Jackfruit Bread

Appearance refers to visual qualities such as colour, shape, and crumb structure. The results showed significant differences among the three samples (p = 0.00), with the control sample (100% wheat flour) receiving the highest appearance score, followed by 20% jackfruit flour, while 40% substitution scored the lowest (Figure 4). Higher substitution levels led to a darker crumb and less uniform loaf volume. This is consistent with findings in fibre-enriched breads, where gluten dilution reduces gas retention and browning uniformity. Post hoc analysis (Bonferroni) confirmed that all sample pairs differed significantly.

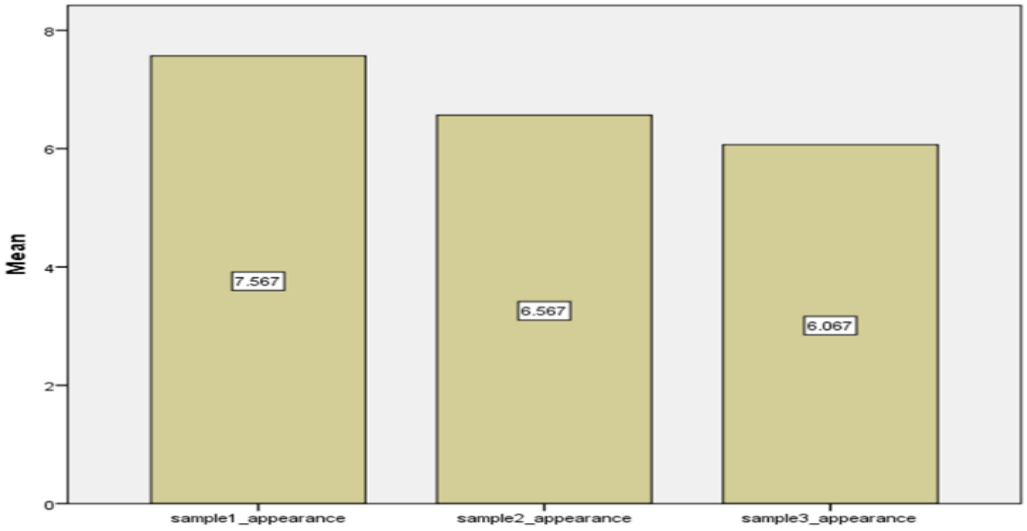


Figure 4: Mean appearance scores of jackfruit bread for the three sample formulations

The study shows that higher substitution levels resulted in darker crumb and less desirable loaf structure due to decreased gluten availability and increased fibre content. This aligns with studies that show non-gluten flours reduce gas retention in doughs (Cappelli et al. 2020) and others that highlight how increased dietary fibre promotes intensified Maillard reaction browning in high-fibre bread products (El Hosry 2025; Schütte 2024).

Aroma of Jackfruit Bread

Aroma is influenced by volatile compounds released during baking. ANOVA showed significant differences (p = 0.00). The control sample had the highest aroma score, followed by 20% jackfruit flour. The 40% sample received the lowest score, likely due to the stronger natural aroma of jackfruit flour (Figure 5).

Research indicates that jackfruit flour contains distinct volatile esters that may overpower typical bread aroma at higher substitution levels. Post hoc results showed all samples differed significantly from one another.

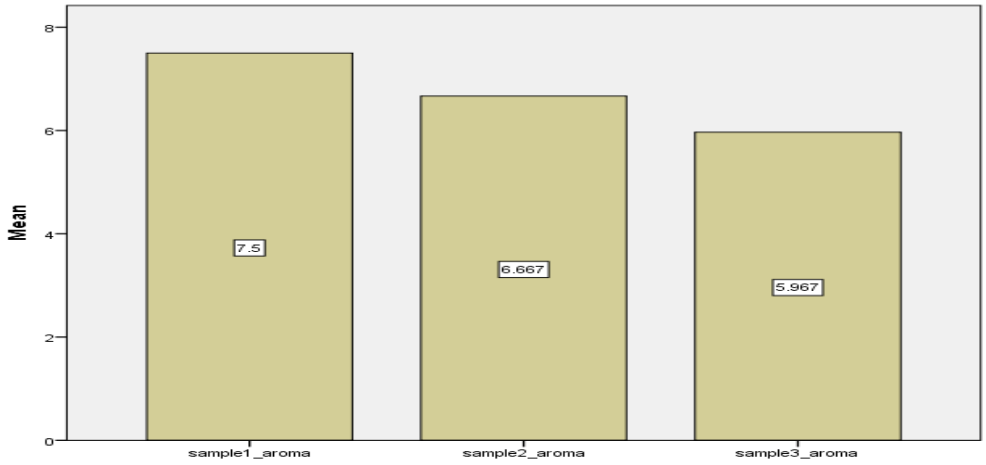


Figure 5: Mean aroma scores of jackfruit bread for the three sample formulations

Jackfruit's natural volatile compounds produced a stronger, fruit-like aroma in the bread, which became particularly pronounced at the 40% substitution level. This is consistent with findings in recent literature, which report that tropical fruit–based flours contribute distinct aromatic profiles due to their ester, aldehyde, and ketone compounds. However, these intensified aromas may not always align with consumer expectations for conventional wheat bread, making them less acceptable at higher substitution levels (Bharath et al., 2023; Nunes et al., 2024).

Texture of Jackfruit Bread

Texture is one of the most critical sensory attributes, encompassing softness, chewiness, and mouthfeel. Significant differences were found across samples (p = 0.00). The control sample scored the highest, followed by 20% substitution, while 40% substitution had the lowest score (Figure 6). This trend is expected because jackfruit flour contains no gluten. Increasing the flour ratio reduces dough elasticity, loaf volume, and crumb softness—common characteristics of gluten-reduced bread. Post hoc results confirmed all samples were significantly different.

The reduction in texture scores is closely linked to the absence of gluten in jackfruit flour. As substitution levels increase, the dough loses its ability to form a cohesive gluten network, resulting in reduced elasticity, increased firmness, and a denser crumb structure. These textural changes are widely reported in current gluten-free and composite-flour bread research, which consistently shows that non-gluten flours weaken gas retention, limit dough expansion, and produce tighter crumb formations (Cappelli et al., 2020; Sciarini et al., 2024).

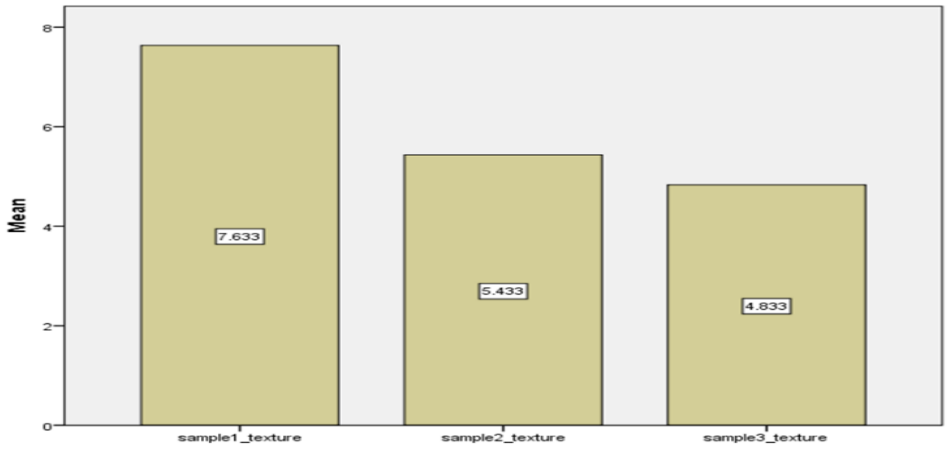


Figure 6: Mean texture scores of jackfruit bread for the three sample Formulations Taste of Jackfruit Bread

Taste significantly decreased with higher substitution (p = 0.00). The familiar wheat-based flavour in the control sample was preferred. The 20% jackfruit bread remained acceptable, while 40% substitution was least preferred (Figure 7). Higher jackfruit flour content introduces a stronger fruit-based flavour, which some panelists

described as unfamiliar in bread. All sample differences were statistically significant based on Bonferroni tests.

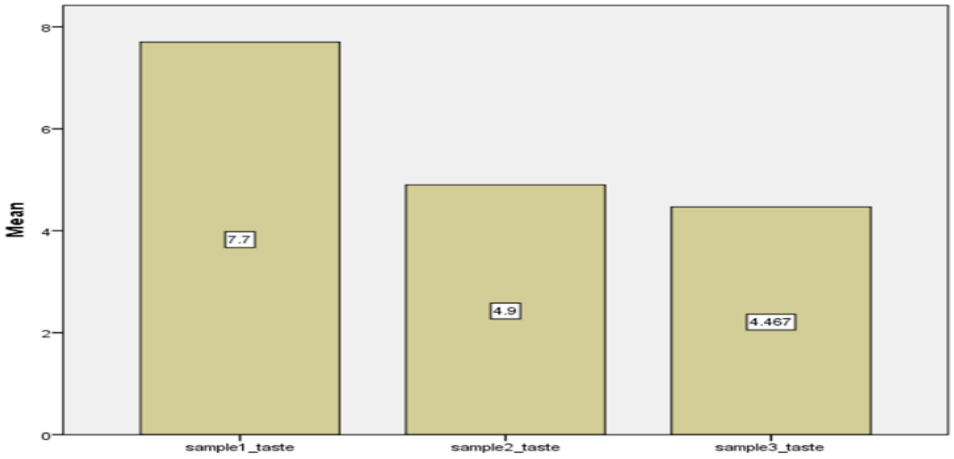


Figure 7: Mean taste scores of jackfruit bread for the three sample formulations

Taste remained acceptable at the 20% substitution level but declined sharply at 40%, indicating a sensory threshold at which jackfruit flour begins to dominate the flavour profile. At higher levels, the naturally sweet, fruity, and slightly musky notes of jackfruit become more pronounced, altering the expected taste of conventional bread. Similar findings have been reported in recent studies, which show that fruit-based and high-fibre flours can introduce strong characteristic flavours that may reduce consumer acceptance when added beyond moderate substitution levels (Santos et al., 2023; Devi et al., 2024).

Overall Acceptance of Jackfruit Bread

Overall acceptance integrates all sensory attributes. Consistent with previous results, the control sample had the highest score, followed by 20% substitution, and finally 40% substitution (Figure 8). This indicates that up to 20% jackfruit flour can be incorporated while maintaining acceptable consumer preference across attributes. Post hoc analysis confirmed significant differences among all sample comparisons.

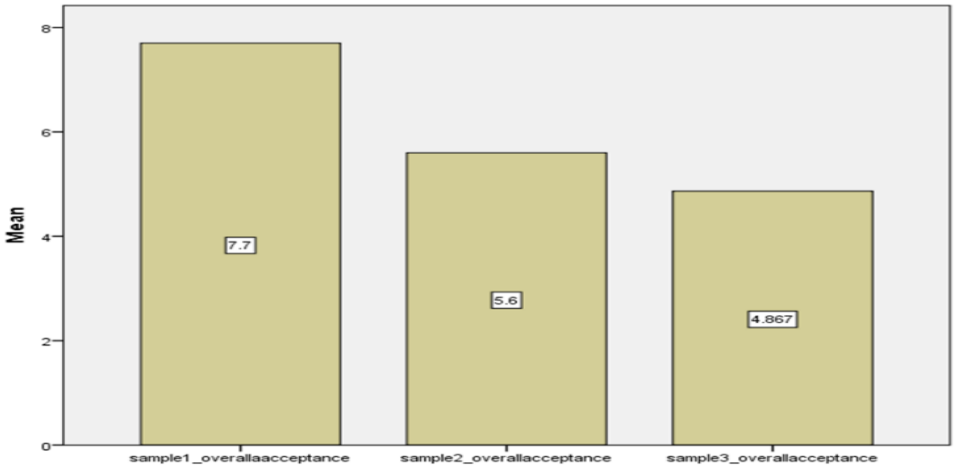


Figure 8: Mean overall acceptance scores of jackfruit bread for the three sample formulations

Overall acceptability indicates that 20% jackfruit flour is the optimal substitution level, as it enhances the bread's nutritional value while maintaining acceptable sensory quality. At this level, the product retains desirable characteristics in appearance, aroma, texture, and taste, without introducing dominant flavours or structural changes that negatively affect consumer preference. In contrast, the 40% substitution level shows a marked decline in acceptability, as the negative effects on texture, aroma, and taste outweigh the potential nutritional

benefits. These findings are consistent with recent studies reporting that moderate incorporation of fruit-based or high-fibre flours can be well tolerated in baked products, but higher levels are often disliked due to stronger flavours and compromised crumb structure (Khan et al., 2023; Velásquez et al., 2024).

Comparative Sensory Profile

A combined radar chart was generated for visual comparison. The radar chart illustrates the mean sensory scores for appearance, aroma, texture, taste, and overall acceptability for the three jackfruit bread samples. The results show a clear decline in sensory acceptance as the level of jackfruit flour substitution increases (Figure 9).

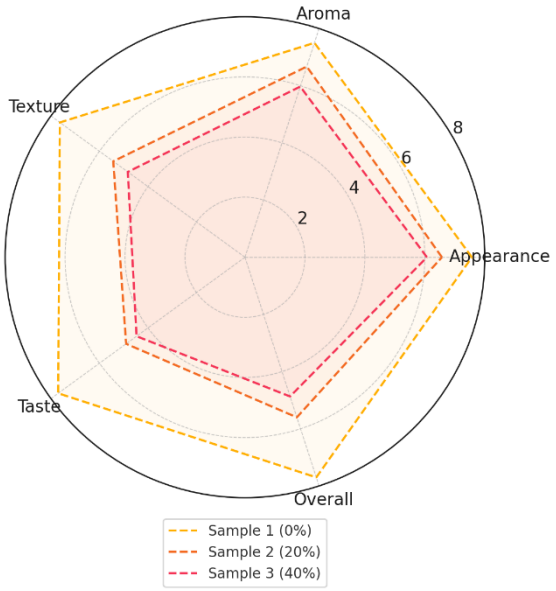


Figure 9: Radar chart of the mean sensory scores for the three jackfruit bread samples

Sample 1 (0% jackfruit flour) recorded the highest scores across all attributes, indicating that the control bread was the most preferred by the panelists. Sample 2 (20% substitution) showed moderate scores, suggesting that partial replacement of wheat flour with jackfruit flour is still acceptable in terms of sensory quality. Sample 3 (40% substitution) received the lowest ratings for all parameters, indicating that a higher proportion of jackfruit flour negatively affected the bread's texture, taste, and overall acceptability. Overall, the chart demonstrates that while small amounts of jackfruit flour can be used without major sensory drawbacks, higher substitution levels reduce consumer acceptance of the bread. It shows a gradual decline across all attributes as substitution increases. The 20% sample demonstrates a moderate reduction from the control but remains within an acceptable sensory threshold.

CONCLUSION

This study examined the effects of substituting wheat flour with jackfruit (Artocarpus heterophyllus) flour at 20% and 40% levels on the sensory properties of bread. The results demonstrated that jackfruit flour significantly influenced appearance, aroma, texture, taste, and overall acceptability across all substitution levels. Bread made with 20% jackfruit flour maintained acceptable sensory quality and scored moderately across all evaluated attributes, indicating that this level of substitution offers a viable balance between nutritional enhancement and consumer preference.

In contrast, the 40% substitution level led to substantial declines in sensory acceptance, particularly in texture and taste, largely due to gluten dilution, increased crumb density, and stronger jackfruit-derived flavour and aroma. These findings suggest that while jackfruit flour provides nutritional advantages such as higher dietary fibre and lower glycaemic potential, its use in bread should be limited to moderate levels to preserve sensory appeal.

Overall, this study supports the incorporation of jackfruit flour as a functional ingredient in bakery applications, especially at substitution levels up to 20%. Future research is recommended to explore textural optimisation techniques, enzymatic improvers, hydrocolloid incorporation, and instrumental analyses to enhance product quality and broaden the commercial feasibility of jackfruit-based functional bread.

ACKNOWLEDGEMENT

This research was not funded by any grant.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author(s) used ChatGPT 5.1 to draft text and format citations. The author(s) subsequently reviewed, revised, and approved all content, and accept full responsibility for the final manuscript.

REFERENCES

Augustin, M. A., Riley, M., Stockmann, R., Bennett, L., Kahl, A., Lockett, T., Osmond, M., Sanguansri, P., Stonehouse, W., & Cobiac, L. (2015). Role of food processing in food and nutrition security. Trends in Food Science & Technology, 47(2), 115–125. https://doi.org/10.1016/j.tifs.2015.11.003

Bharath, K. N., et al. (2023). Influence of tropical fruit powders on aroma and sensory quality of baked products. Journal of Food Processing and Preservation.

Cappelli, A. (2020). A systematic review of gluten-free dough and bread. Applied Sciences, 10(18).

Cappelli, A., et al. (2020). A systematic review of gluten-free dough and bread. Applied Sciences, 10(18).

Costa, A., Dinnella, C., Miele, N., & Monteleone, E. (2023). Advances in sensory evaluation methods for functional and health-oriented foods. Food Quality and Preference, 109, 104902. https://doi.org/10.1016/j.foodqual.2022.104902

Devi, R., et al. (2024). Effect of fruit powder incorporation on flavour and acceptability in composite-flour baked products. Food Research International.

El Hosry, L. (2025). Maillard Reaction: Mechanism, influencing parameters... Foods, 14(11), 1881.

Engindeniz, S., & Bolatova, K. (2019). Global consumption patterns and nutritional importance of bread. Journal of Food and Nutrition Research, 58(1), 20–28.

Fasano, A., Sapone, A., Zevallos, V., & Schuppan, D. (2015). Nonceliac gluten sensitivity. Gastroenterology, 148(6), 1195–1204. https://doi.org/10.1053/j.gastro.2014.12.049

Ferrari, L., Bonomi, F., Iametti, S., & Pagani, M. A. (2022). Development of low-glycaemic index bread using alternative cereal and tuber flours. Food Research International, 157, 111304. https://doi.org/10.1016/j.foodres.2022.111304

Ghosh, A., & Chattopadhyay, P. (2021). Effect of fiber-rich composite flours on bread quality. Journal of Cereal Science, 98, 103150. https://doi.org/10.1016/j.jcs.2021.103150

IDF. (2021). Global diabetes overview. International Diabetes Federation Atlas.

Joseph, B., Rabindran, R., & George, A. (2018). Nutritional and functional properties of jackfruit flour. Food Chemistry, 245, 123–128.

Khan, M. A., et al. (2023). Consumer acceptability of composite-flour breads enriched with fruit and fibre sources. Journal of Food Quality.

Kumar, R., Das, A., & Thomas, P. (2023). Emerging tropical fruit flours in functional food applications: A review. Current Nutrition & Food Science, 19(2), 123–134. https://doi.org/10.2174/1573401318666210811154121

Melini, V., & Melini, F. (2021). Strategies for developing gluten-free bakery products: Recent advances and future perspectives. Foods, 10(10), 2359. https://doi.org/10.3390/foods10102359

Nunes, C. A., et al. (2024). Volatile compound profiles in fruit-based flours and their impact on baked product acceptability. Food Chemistry.

Patel, J., Singh, A., & Rao, M. (2021). Influence of jackfruit flour on dough rheology and bread characteristics. Journal of Food Processing and Preservation, 45, e15890.

Reddy, N., Venkatesh, K., & Rao, D. (2022). Nutritional composition and functional properties of unripe jackfruit flour for food applications. Food Chemistry, 373, 131464. https://doi.org/10.1016/j.foodchem.2021.131464

Schütte, M. (2024). Nonenzymatic browning and antioxidant properties of whole wheat bread. Journal of Food Science & Technology.

Sciarini, L. S., et al. (2024). Impact of composite flours on dough rheology and gluten-replacement strategies in breadmaking. Food Hydrocolloids.

Shelar, A., Khedkar, R., & Manohar, R. S. (2023). Utilisation of underused tropical fruits in bakery products: Technological and nutritional perspectives. Journal of

Food Processing and Preservation, 47(1), e17242. https://doi.org/10.1111/jfpp.17242

Shilpa, S., & Lakshmi, B. (2021). Development of functional bakery products using jackfruit flour: Nutritional and sensory properties. Journal of Food Science and Technology, 58, 2501–2509. https://doi.org/10.1007/s13197-020-04707-0

Santos, J. P., et al. (2023). Sensory thresholds of fruit-enriched bread formulations and their impact on consumer liking. Journal of Cereal Science.

Velásquez, B., et al. (2024). Impact of high-fibre flour substitution on sensory attributes and overall liking in bread products. Food Chemistry.

Zhang, Y., Li, X., & Wang, H. (2019). Effect of gluten dilution and fibre enrichment on bread texture and structure. Food Hydrocolloids, 87, 420–428. https://doi.org/10.1016/j.foodhyd.2018.08.045

ORIGINAL ARTICLE

A Survey on The Acceptance of Non-Nutritive Sweeteners among Diabetic Patients

Nur Wadhihah Abdul Jalil[1], Noor Suzana Osman[1,2], Roszanadia Rusali [1,2], Wan Fathin Fariza Wan Mahmood[1,2,*]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
non-nutritive sweeteners; stevia; erythritol; diabetes, sensory evaluation

ABSTRACT

Background: The prevalence of Type 2 Diabetes Mellitus (T2DM) in Malaysia continues to rise, largely due to excessive sugar consumption. Non-nutritive sweeteners (NNS), such as stevia and erythritol, are increasingly used as alternatives due to their low caloric content and minimal impact on blood glucose. However, the acceptance of NNS among diabetic patients remains uncertain. This study aimed to assess the acceptance of a stevia-erythritol blend among hospitalised T2DM patients. **Methods:** A cross-sectional sensory evaluation was conducted among 69 T2DM patients hospitalised at one of the teaching hospitals in East Coast Malaysia. Participants consumed tea sweetened with a stevia-erythritol blend and rated its sweetness, aftertaste, and overall acceptability using a 5-point hedonic scale. Demographic data, body mass index (BMI), blood glucose levels, and length of hospital stay were also recorded. **Results:** Most of the participants (66.7%) accepted the NNS-sweetened tea. Mean scores indicated positive responses for sweetness (3.71 ± 0.86), aftertaste (3.70 ± 0.85), and overall acceptability (3.86 ± 0.91). **Conclusion:** The findings indicate that diabetic patients generally accepted NNS, with sensory attributes such as sweetness and aftertaste playing a key role in their acceptance. These results support the inclusion of more palatable NNS-based options in hospital food services to promote better dietary adherence and glycaemic control in T2DM management.

INTRODUCTION

According to the International Diabetes Federation (IDF), approximately 537 million adults aged 20-79 are living with diabetes, and this number is expected to rise to 643 million by 2030 and 783 million by 2045. The IDF predicts that by 2045, the number of people with diabetes in the Southeast Asia (SEA) Region will increase by 68%, reaching 152 million, with a diabetes prevalence of 11.3%. In Malaysia specifically, the prevalence of diabetes is particularly concerning, as it has been reported that approximately over half a million, or 2.5% of adults in Malaysia live with four non-communicable diseases (NCDs), which are diabetes, hypertension, high cholesterol, and obesity.

The rising prevalence of T2DM is concerning due to its often subtle symptoms, which delay early detection**.** The symptoms of T2DM take years to develop and become noticeable, unlike type 1 diabetes, where the symptoms appear more obvious. The Ministry of Health Malaysia (2012) emphasizes that the key strategy for achieving glycaemic control is to monitor total carbohydrate intake, including sugar intake.

Recent research has proved that excessive sugar consumption is associated with the increasing trend in the prevalence of T2DM (Liu et al., 2023). According to the World Health Organization (WHO) guidelines on sugar intake for adults and children, it is strongly recommended to reduce the intake of free sugar throughout life. The WHO also recommends reducing free sugar intake to less than 10% of total energy, with further reductions to less than 5% for additional health benefits (WHO, 2015).

Health concerns over excessive sugar intake have increased NNS use, especially among individuals with obesity and diabetes. (Lohner et al., 2020). NNS offer sweetness without significant calories and are used as an alternative to nutritive sweeteners to help lower energy intake and minimize carbohydrate-related health risks. Non-nutritive sweeteners are also often recommended for individuals with T2DM due to their minimal impact on glycaemic levels (Ministry of Health Malaysia, 2012). These sweeteners can serve as an alternative to regular sugar, helping diabetic patients better manage their blood glucose levels. However, it is essential to ensure that their intake does not exceed the Acceptable Daily Intake (ADI) levels as recommended by the Food and Drug Administration (FDA) (Ministry of Health Malaysia, 2012).

[1]* *Corresponding author.*
E-mail address: wfathin@iium.edu.my

In Malaysia, the use of stevia as a sugar alternative is increasingly promoted, particularly in response to rising health concerns like diabetes and obesity (Saharudin et al., 2020). Based on the study among health-educated individuals, a significant majority agreed that stevia could substitute sugar and would recommend it for diabetic patients (Saharudin et al., 2020). However, the aftertaste of stevia affects user acceptance because it leaves a lingering unpleasant taste. Erythritol, which is nearly 70% as sweet as sucrose, has a mild cooling effect in the mouth with no aftertaste (Mazi & Stanhope, 2023). It can increase volume and reduce the aftertaste associated with stevia, particularly its bitter and licorice-like notes (Myers et al., 2018). Therefore, stevia is often combined with erythritol to enhance the taste and usability of the sweetener blend while minimizing common drawbacks associated with pure stevia. Thus, this study aimed to determine the acceptance of NNS among diabetes patients.

MATERIALS AND METHODS

Participants

A cross-sectional sensory evaluation study was conducted among inpatients with T2DM in Sultan Ahmad Shah Medical Centre (SASMEC @IIUM), aged 18 years and above. The data collection was done from February to May 2025. This study excluded T1DM patients, pregnant diabetic patients, patients on fluid restriction, patients in isolation, tube feeding patients, and patients with cognitive impairments.

Sampling Strategy

Purposive sampling was used to recruit 105 T2DM patients, who had been hospitalised at SASMEC @IIUM and were willing to participate. Participants could withdraw at any time. Written consent was obtained from each participant before their involvement in the study. The consent form, survey form, and participant agreement were provided and required before the assessments were conducted. A list of patients receiving a diabetic diet was obtained from the hospital food service dietitians. From this list, patients with fluid restrictions and those in isolation rooms were excluded from the study. The researchers then approached the remaining eligible patients to obtain their consent to participate.

Data Collection

In this study, the data collection form was prepared in the Malay language to ease communication with local participants. The form included two components which were 1) participants' background information that consisted of demographic (age and gender), anthropometry (weight and height), blood glucose level (capillary blood glucose), and length of hospital stay, and 2) sensory evaluation that focused on aftertaste, sweet taste, and overall acceptability of the tea with NNS.

Participants' Background Information

During the data collection process, participants were asked to complete a survey form that included background information on their hospital registration number (RN), age, gender, weight, height, and the number of days they had been hospitalised at the time they consumed the tea. Their registration number were used to access the hospital's electronic medical record system. For participants who were unable to recall or provide their weight and height, the data were obtained from their medical records.

These measurements were then used by the researcher to calculate each participant's Body Mass Index (BMI). Then the value of BMI was categorized according to Malaysian Clinical Practice Guidelines for Obesity (2023) BMI based weight classification where underweight (less than 18.5 kg/m^2), normal weight (18.5-22.9 kg/m^2), overweight (23-27.4 kg/m^2), and obese (27.5 kg/m^2 or greater) for participants ages less than 60 years old. For participants aged more than 60 years old, their BMI was classified according to Malaysian Dietary Guidelines for Older Persons (2023)' BMI classification, where 24 to 27 kg/m^2 indicates the normal range, while below and above the range indicate underweight and obese.

According to the Malaysian Clinical Practice Guidelines for T2DM (2020), the recommended 2-hour post-prandial blood glucose target range for patients with Type 2 diabetes is between 4.4 mmol/L and 8.5 mmol/L. Therefore, the reading between 4.4 mmol/L and 8.5 mmol/L is considered within the target range, and higher and lower indicate below and above the target range, respectively.

The measurement for blood glucose was capillary blood glucose, and all data were retrieved from the medical record to ensure consistency. All values were recorded between 12:00 p.m. to 1:00 p.m., before the tea was consumed. Following participants' discharge, the total length of hospital stay was retrieved from the medical record.

Sensory Evaluation

In this study, a sensory evaluation form was used to assess participants' responses to the tea containing non-nutritive sweeteners (NNS). This method directly measured

participants' preferences and satisfaction using a 5-point hedonic scale (Addo-Preko et al., 2023). The 5-point hedonic scale consisted of five levels of liking, ranging from 1 = "dislike very much" to 5 = "like very much," with 3 = "neither like nor dislike" representing the neutral point.

Table 1: Description for Sensory Evaluation Score

Score	Description
1	Dislike Extremely
2	Dislike Slightly
3	Neither Like nor Dislike
4	Like Slightly
5	Like Extremely

While some studies have defined a mean score of 4 or higher as indicating strong acceptability (Everitt, 2009), other researchers interpret mean scores above 3.0 as reflecting positive consumer acceptance, particularly when converted to an acceptance index exceeding 70% (Huey et al., 2023). In this study, the sensory attributes evaluated included aftertaste, sweetness, and overall acceptability. Therefore, in this study, the acceptance index approach was applied, and the results were interpreted based on whether the mean scores indicated an acceptable level of consumer preference. The sensory attributes evaluated included aftertaste, sweetness, and overall acceptability.

$$\text{Acceptance Index (AI)} = \frac{Mean\ Hedonic\ Score}{Maximum\ Possible\ Score} \times 100$$

A standard sachet of NNS weighs 2.5 g and is designed to provide sweetness equivalent to one teaspoon of sugar (EviaSweet, Fine Foods, Malaysia). It is a combination of erythritol and stevia (proportion not disclosed).

Tea was prepared by immersing a tea bag (Cameron Valley black tea, Malaysia) in 200 ml of boiling water for one serving. Tea was steeped for 3 minutes, allowing the leaves to release their full bouquet of flavor and aroma. The tea bag was then taken out of the drink to avoid overpowering the taste. Then, one sachet of NNS (2.5 g) was added to the tea.

Research Flow

The tea sample preparation was carried out by the researcher in the hospital kitchen around 2:00 pm. The tea was then served during afternoon tea between 3.00 p.m. to 4.00 p.m., along with the data collection form. The completed forms were then collected around 4:30 pm.

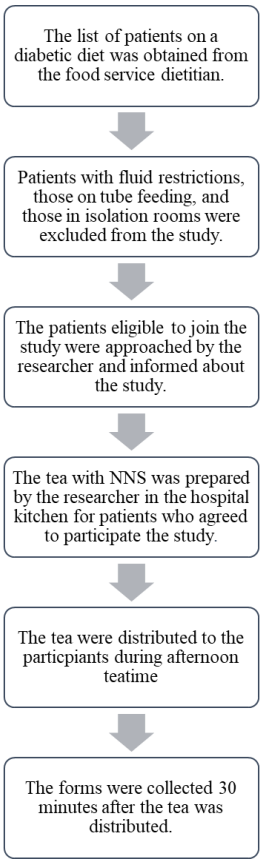


Figure 2: Participants' recruitment process and sensory evaluation protocol

Ethical Approval

This study received ethical approval from the Kulliyyah Postgraduate and Research Committee (KPGRC) (KAHS 17/25), the International Islamic University Malaysia Research Ethical Committee (IREC) (IREC 2025-KAHS/DNS13), and the Department of Education and Research, Sultan Ahmad Shah Medical Centre @IIUM (IIR25-16).

RESULTS

There was a total of 69 participants who participated in this study. Despite 155 patients in SASMEC@IIUM being approached to be participants in this study, only 69 patients agreed to be participants, which resulted in a low response rate (44.5%). However, only 49 participants with complete information on their anthropometric measurements.

Characteristics of the Participants

Table 2 demonstrates the sociodemographic characteristics of the total participants (N=69), with 60.9%

being female (n=42) and 39.1% being male (n=27). The age group of 60-69 years old comprised the largest percentage of participants (39.1%, n=27), while those 39 years old and younger represent the lowest proportion (2.9%, n=2). The ethnic composition of the participants was predominantly Malay (95.7%, n= 66), and only 3 participants were of a different race, which were Chinese (1.4%, n=1) and Indian (2.9%, n=2). In terms of length of hospital stay, most participants were hospitalised for 1-7 days (47.8%, n=33), whereas the longest recorded length of stay was 40 days (1.4%, n=1)

BMI Status of Participants

The anthropometric measurements of the 49 participants are presented in Table 3. The mean body mass index (BMI) was 27.35±5.71 kg/m^2. According to Malaysian Clinical Practice Guidelines for Obesity (2023) and Malaysian Dietary Guidelines for Older Persons (2023), most participants, 51.0% (N=25), were obese, followed by 24.5% (N=12) normal, 22.4% (N=11) underweight, and 1.4% (N=1) overweight. Obesity was more predominant among females (53.8%) than males (47.8%). A higher proportion of males (30.4%) were in the normal weight range compared to females (19.2%).

Table 2: Sociodemographic Characteristics of Participants (N=69)

Variables	n (%)	Mean ± SD
Age		64.74 ± 9.61
30-39 years	2 (2.9)	
40-49 years	3 (4.3)	
50-59 years	12 (17.4)	
60-69 years	27 (39.1)	
70-79 years	21 (30.4)	
≥80 years old	4 (5.8)	
Gender		
Male	27 (60.9)	
Female	42 (39.1)	
Race		
Malay	66 (95.7)	
Chinese	1 (2.9)	
Indian	2 (1.4)	
Length of Hospital Stay		10.32 ± 7.77
1-7 days	33 (47.8)	
8-14 days	20 (29.0)	
15-21 days	9 (13.0)	
22-28 days	5 (7.2)	
29-35 days	1 (1.4)	
36-42 days	1 (1.4)	

Blood Glucose Level

Capillary blood glucose readings were obtained from all 69 participants to assess glucose control during hospitalization. In this study, the mean capillary blood glucose was 8.17 ± 3.06 mmol/L. Among the participants, 3 (4.3%) had glucose levels below the target range, 43 (62.3%) were within the recommended target, and 23 (33.3%) had above-target glucose levels, indicating that a significant proportion of participants experienced suboptimal glucose control during hospitalization. When comparing glucose control by gender, the distribution was generally similar. There were more male and female participants with normal blood glucose control (61.9% among males, 63% among females). Notably, participants with below-target glucose levels were all female (7.1% of all females). However, these findings suggest no apparent difference in glucose control status between genders in this study group.

Sensory evaluation of tea with non-nutritive sweeteners (NNS)

Descriptive statistics from the 5-point hedonic scale showed that, for sweetness, the majority of participants (74%) neither like nor dislike it, with 21.7% liking it extremely and 31.9% slightly liking it. The mean sweetness score was 3.71 ± 0.86, corresponding to an acceptance index (AI) of 74.2%, indicating that the product was within the acceptable range of sensory liking (AI ≥ 70%)
A similar pattern was observed for aftertaste, where 20.3% of participants liked it extremely, and 33.3% slightly liked it, resulting in a mean score of 3.70 ± 0.85 and an AI of 74.0%, also reflecting a generally acceptable level of sensory response.

The overall acceptability of the tea with NNS was the highest among the evaluated attributes. Most participants (40.6%) slightly liked it, and 26.1% liked it extremely, with only 1.4% expressing extreme dislike. The mean overall acceptance score of 3.86 ± 0.91 corresponded to an AI of 77.2%, suggesting that participants were generally receptive to and satisfied with the taste and presence of non-nutritive sweeteners in the tea samples.

Table 3: Anthropometric assessment and BMI categories of participants (N=49)

Variables	n (%)	Mean ± SD
BMI (kg/m^2)		27.35 ± 5.71
Underweight	**11 (22.4)**	
Male	5 (21.7)	
Female	6 (23.1)	
Normal	**12 (24.5)**	
Male	7 (30.4)	
Female	5 (19.2)	
Overweight	**1 (1.4)**	
Male	0	
Female	1 (3.8)	
Obese	**25 (51.0)**	
Male	11 (47.8)	
Female	14 (53.8)	

Table 4: Capillary Blood Glucose Levels of Participants (N=69)

Variables	n (%)	Mean ± SD
Capillary blood glucose (mmol/L)		8.17 ± 3.06
Below Target <4.4 mmol/L)	**3 (4.3)**	
Male	0 (0.0)	
Female	3 (7.1)	
Within Target (4.4 – 8.5 mmol/L)	**43 (62.3)**	
Male	17 (61.9)	
Female	26 (63.0)	
Above Target (> 8.5 mmol/L)	**23 (33.3)**	
Male	10 (37.0)	
Female	13 (31.0)	

Table 5: Mean hedonic scores for sensory evaluation taste of NNS (N=69)

Variables	n (%)	Mean ± SD	Acceptance Index (AI) (%)
Sweetness		3.71 ± 0.86	74.2
Dislike Extremely	0 (0)		
Dislike Slightly	3 (4.3)		
Neither like nor Dislike	29 (42.0)		
Like Slightly	22 (31.9)		
Like Extremely	15 (21.7)		
Aftertaste		3.70 ± 0.85	74.0
Dislike Extremely	0 (0)		
Dislike Slightly	3 (4.3)		
Neither like nor Dislike	29 (42.0)		
Like Slightly	23 (33.3)		
Like Extremely	14 (20.3)		
Overall Acceptance		3.86 ± 0.91	77.2
Dislike Extremely	1 (1.4)		
Dislike Slightly	3 (4.3)		
Neither like nor Dislike	19 (27.5)		
Like Slightly	28 (40.6)		
Like Extremely	18 (26.1)		

DISCUSSION

This study aims to evaluate the acceptance of non-nutritive sweeteners (NNS), specifically the combination of stevia and erythritol, among hospitalised diabetic patients. Conducted with a sample of 69 patients, the findings provide valuable insights into NNS acceptance within this population. Recent studies have highlighted key factors influencing the acceptance of non-nutritive sweeteners (NNS), particularly stevia and erythritol, among various consumer groups, including individuals with diabetes (Hong et al., 2025; Mazi & Stanhope, 2023). While stevia is widely recognized for its health benefits, especially its non-glycaemic properties, its sensory qualities often present a barrier to acceptance. The most frequently reported concern is its lingering aftertaste, which many users find bitter or metallic. This aftertaste has been consistently cited as a factor that detracts from the overall sensory experience, especially when stevia is used as a standalone sweetener (Saharudin et al., 2020).

To improve consumer satisfaction, erythritol is often blended with stevia. Erythritol, a sugar alcohol with a clean, sweet taste and minimal aftertaste, helps to balance Stevia's more pronounced sensory attributes. According to Pourahmad and Khorramzadeh (2016), combining erythritol with stevia significantly improved the overall sensory acceptance of a soy-based beverage powder. Their study found that the formulation using 80% stevia and a

75:25 erythritol-to-isomalt ratio received the highest scores for sweetness, mouthfeel, sweetness aftertaste, and overall acceptability. The use of erythritol was shown to reduce the undesirable aftertaste often associated with stevia, resulting in a sweetness profile that more closely resembled that of sucrose.

These results are consistent with the findings of this study, which indicate a favourable perception of NNS among the participants in terms of sweetness, aftertaste, and overall acceptance. For sweetness, most participants expressed a positive or neutral response. The mean sweetness score indicated a moderate level of liking. Based on the calculated Acceptance Index (AI), sweetness (74.2%), aftertaste (74.0%), and overall acceptance (77.2%) all exceeded the 70% threshold typically considered indicative of acceptable sensory quality (Huey et al., 2023), supporting that the NNS used in this study was well accepted by participant The overall acceptance of non-nutritive sweeteners (NNS) was the highest among the evaluated attributes, indicating that most hospitalised diabetic patients were receptive to their taste. A similar positive response was seen for sweetness and aftertaste. The strong and significant associations between these taste attributes and overall acceptance suggest that the palatability of NNS, particularly sweetness and aftertaste, plays a key role in influencing their acceptance in this population.

Demographic factors such as age, gender, and BMI did not show a significant association with NNS acceptance. This indicates that sensory experience. especially taste and aftertaste, was more decisive than individual characteristics (Farhat et al., 2021; Moreira et al., 2025). The predominance of older adults in this study may have positively influenced acceptance, as previous research has shown that older Malaysians generally favor natural, plant-based products (Kim Sooi & Lean Keng, 2013). Given that stevia is a plant-derived sweetener, this demographic's preference for natural health alternatives may have enhanced their openness toward the stevia–erythritol blend.

When compared with a previous study conducted at Universiti Kuala Lumpur Royal College of Medicine Perak, which specifically evaluated the use of pure stevia, our findings indicate a notably higher level of acceptance among hospitalised diabetic patients. In the earlier study, while a significant proportion of participants (66.7%) liked the taste of stevia and 95.1% were willing to recommend it for diabetic patients, a substantial number (74.1%) reported an unpleasant aftertaste, and 65.4% perceived a noticeable change in the taste of their drinks (Saharudin et al., 2020). In contrast, our study reported a moderately positive perception, with the aftertaste receiving an average score approaching 3.7 and overall acceptance nearing 3.9 on the hedonic scale. Notably, only a minimal proportion of participants, approximately 1%, expressed strong disapproval. This suggests that the non-nutritive sweetener used in our study, stevia-erythritol blend, provided a more favourable sensory experience. The combination of stevia and erythritol has been shown to effectively reduce the bitterness and lingering metallic aftertaste typically associated with pure stevia, resulting in a sweeter taste profile (Mazi & Stanhope, 2023; Pourahmad & Khorramzadeh, 2016). Therefore, the improved acceptance observed in our study may be attributed to the enhanced palatability of this blend, reinforcing its potential as a suitable and well-tolerated sugar substitute for diabetic individuals in clinical settings.

Taken together, these findings suggest that while NNS like stevia and erythritol may not perfectly replicate the taste of sugar, they are generally well-accepted when appropriately formulated. Our results align with the broader literature in indicating that both the sweetness level and aftertaste are significantly associated with overall acceptance. Thus, given the strong acceptance observed, this NNS has the potential to support dietary adherence while enhancing the palatability of low-sugar diets (Huey et al., 2023).

Stevia, while popular for its zero-calorie benefits and glycaemic neutrality, is often criticized for its lingering bitter or metallic aftertaste. Erythritol, on the other hand, offers a sweetness profile and can mask the aftertaste of stevia effectively. By combining these two sweeteners, the study assessed a more consumer-acceptable alternative that more closely mimics the sensory qualities of sugar. This is especially relevant for hospital food services, where taste acceptability plays a crucial role in ensuring patients' compliance with dietary prescriptions.

Beyond sensory perception, health awareness and motivation also appear to influence the acceptance of NNS among participants. Many hospitalised diabetic patients in this study were already familiar with sugar restriction as part of their dietary management, which may enhance their openness toward NNS as a healthier sugar substitute. This aligns with the recommendations of the American Diabetes Association (ADA, 2022), which emphasize the use of low- or non-caloric sweeteners to aid glycaemic control and support dietary compliance. The finding also supports the WHO's guideline stating that NNS can be used as a tool to reduce free sugar intake, especially among populations at risk of non-communicable diseases such as diabetes (WHO, 2023). This suggests that participants' cognitive acceptance, derived from awareness of health

benefits, may have reinforced their sensory acceptance of the stevia-erythritol blend. This is consistent with previous findings that diabetic individuals who are aware of the health implications of sugar intake tend to show greater acceptance of NNS as part of dietary management (Farhat et al., 2021).

Overall, these findings emphasize that incorporating acceptable NNS formulations into hospital food services may serve as a practical dietary strategy to enhance compliance without compromising taste or nutritional goals. Conducting this study within a real hospital environment adds clinical relevance, as it reflects the actual context of diabetic dietary management. Such an approach not only supports dietary adherence but also aligns with public health recommendations promoting the reduction of added sugar intake for chronic disease prevention and management.

Strengths and Limitations

This study on the acceptance of NNS, specifically a blend of stevia and erythritol, among hospitalised diabetic patients presents several strengths that enhance the credibility and practical relevance of the findings. By focusing on hospitalised diabetic patients, the study provides insights that are highly applicable to hospital-based dietary practices and nutritional care. This translational health research allowed for the assessment of NNS acceptance in a realistic context where patients were already engaged in medical nutrition therapy, thereby increasing the ecological validity of the research. Another key strength is the use of a stevia-erythritol blend rather than relying on single sweeteners.

For further research, it is also important to recognize that external variables among the participants, such as current medication, illness severity, taste alterations due to malnutrition, and psychological factors, could have influenced the sensory evaluation results. Many hospitalised patients experience altered taste perception due to their medical condition or treatment regimen, which may have impacted their ability to accurately assess the sweetness or aftertaste of the NNS.

CONCLUSION

This study indicates that NNS, specifically a blend of stevia and erythritol, is generally acceptable among hospitalised diabetic patients. The majority of participants rated the tea containing NNS positively in terms of sweetness, aftertaste, and overall acceptance. This indicates that NNS can serve as a suitable alternative to sugar in diabetic diets without compromising taste satisfaction.

These results can be used to inform hospital food service practices, particularly in designing diabetic-friendly meals. The positive reception of the stevia-erythritol blend supports its potential inclusion in hospital menus as a healthier alternative to sugar, helping improve the palatability of meals and encouraging better dietary compliance among diabetic patients.

In conclusion, this study provides valuable evidence supporting the use of blended nonnutritive sweeteners as a practical strategy to enhance hospital meals and support effective dietary management in patients with T2DM.

Recommendation

Future studies should consider using a broader and more diverse of participants to enhance the generalizability of the results. Including diabetic patients from various hospitals, regions, and cultural backgrounds would provide a more accurate reflection pf overall acceptance of non-nutritive sweeteners.. It is also recommended that future research include a more comprehensive assessment of participants' nutritional status. This will help to identify whether malnutrition plays a role in taste perception, which could affect the way NNS are received and rated during the sensory evaluation.

ACKNOWLEDGEMENT

This research was funded by LCP Supplies (M) Sdn. Bhd. We thank the staff from IRKAZ Gourmet Sdn. Bhd (SASMEC @IIUM branch) for assisting in the data collection processes.

REFERENCES

American Diabetes Association. (2022). Standards of Medical Care in Diabetes—2022 Abridged for Primary Care Providers. Clinical Diabetes, 40(1). https://doi.org/10.2337/cd22-as01

Farhat, G., Dewison, F., & Stevenson, L. (2021). Knowledge and Perceptions of Non-Nutritive Sweeteners Within the UK Adult Population. Nutrients, 13(2), 444. https://doi.org/10.3390/nu13020444

Hong, Q., Huang, Y., Yang, J., Su, L., Dai, Z., Zhao, C. (2025) Food seetetners: Angels or clowns for human health? Cuurent Research in Food Science, 101032. https://doi.org/10.1016/j.crfs.2025.101032.

Huey, S. L., Bhargava, A., Friesen, V. M., Konieczynski, E. M., Krisher, J. T., Mduduzi N N Mbuya, Mehta, N. H., Monterrosa, E., Nyangaresi, A. M., & Mehta, S. (2023).

Sensory acceptability of biofortified foods and food products: a systematic review. Nutrition Reviews. https://doi.org/10.1093/nutrit/nuad100

International Diabetes Federation. (2021). IDF diabetes atlas: 10th edition. Retrieved November 19, 2024, from https://diabetesatlas.org

Kim Sooi L, Lean Keng S. (2013) Herbal Medicines: Malaysian Women's Knowledge and Practice. Evid Based Complement Alternat Med. 2013:438139. doi: 10.1155/2013/438139.

Liu, Y., Cheng, J., Wan, L., & Chen, W. (2023). Associations between total and added sugar intake and diabetes among Chinese adults: The role of body mass index. Nutrients, 15(14), 3274–3274. https://doi.org/10.3390/nu15143274

Lohner, S., Kuellenberg de Gaudry, D., Toews, I., Ferenci, T., & Meerpohl, J. J. (2020). Non-nutritive sweeteners for diabetes mellitus. Cochrane Database of Systematic Reviews. https://doi.org/10.1002/14651858.cd012885.pub2

Mazi, T. A., & Stanhope, K. L. (2023). Erythritol: An in-depth discussion of its potential to be a beneficial dietary component. Nutrients, 15(1), 204. https://doi.org/10.3390/nu15010204

Ministry of Health Malaysia. (2012). Medical Nutrition Therapy Guidelines for Type 2 Diabetes Mellitus.

Moreira, T.K.B., Mendes, F.D., Santos, H., Quinte, G.C., Mill, J.G., Molina, M.D.C.B., Faria, C.P. (2025) Exploring the Relationship Between Nonnutritive Sweeteners and Nutrient Intake: Findings from the ELSA-Brasil Baseline Study. Nutrients. 24;17(11):1778. doi: 10.3390/nu17111778.

Myers, E. A., Duncan, S. E., Wang, A., & Hedrick, V. E. (2018). Identifying an appropriate carrier for nonnutritive sweeteners in metabolic and controlled feeding investigations via sensory evaluation. Journal of Sensory Studies, 34(2), e12488. https://doi.org/10.1111/joss.12488

Pourahmad, R., & Khorramzadeh, D. (2016). Physicochemical and organoleptic properties of drinking powder containing soymilk powder, stevia, isomalt and erythritol. Journal of Food Processing and Preservation, 40(6), 1189–1200. https://doi.org/10.1111/jfpp.12703

Saharudin, A. M. B., Nazri, N. B. M., Hawi, M. H. B., & Mar, S. O. (2020). Acceptance of stevia as a sugar substitute and its determinants among health educated individuals. Current Research in Nutrition and Food Science Journal, 8(1), 226–237. https://doi.org/10.12944/crnfsj.8.1.21

World Health Organization. (2015, March 4). WHO calls on countries to reduce sugars intake among adults and children. World Health Organization. Retrieved November 19, 2024, from https://www.who.int/news/item/04-03-2015-who-calls-on-countries-to-reduce-sugars-intake-among-adults-and-children

WHO. (2023). Use of non-sugar sweeteners: WHO guideline. Www.who.int. https://www.who.int/publications/i/item/9789240073616

ORIGINAL ARTICLE

Energy and Protein Intake and Its Association with Malnutrition Risk among Community-Living Older Adults in Kuantan, Pahang

Siti Nur Amirah Sheikh Hishamuddin[1], Nur Izzati Mohammad Rosli[2] Aliza Haslinda Hamirudin[2,3*]

[1]Faculty of Health, University of Cyberjaya, Persiaran Bestari, Cyber 11, 63000, Cyberjaya, Selangor
[2]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[3]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia

Keywords:
Older adults; Energy intake; Protein intake; Protein sources; Malnutrition risk

ABSTRACT

Background: Inadequate dietary intake among older adults poses a public health concern as it could lead to adverse effects, including malnutrition risk and poor health outcomes. Thus, this study aimed to assess the energy and protein intake of community living older adults in Kuantan, Pahang, and its association with malnutrition risk. Another objective was to identify sources of protein foods commonly consumed in this studied population. **Methods:** A cross-sectional study was conducted among community-living older adults in Kuantan, Pahang. The dietary intake data were obtained from the diet history method. Malnutrition risk was assessed using Mini Nutritional Assessment-Short Form (MNA-SF). Anthropometric measurements of weight and height were performed for body mass index (BMI) calculation. Data was analysed using descriptive statistics, one-sample t-test, and one-way ANOVA. P-value was set at $p<0.05$ as statistically significant. **Results:** A total of seventy-three (n=73) community living older adults participated in this study, with a mean age of 65.88±5.67 years, and a mean BMI was 26.023±4.21 kg/m^2. Findings showed that energy and protein intake of the respondents were significantly lower than the Recommended Nutrient Intake (RNI). The most common protein sources consumed by the respondents were animal-based sources, particularly fish (80.8%). Eating at home was more practiced by respondents than eating out. Advancing age was significantly associated with malnutrition risk. No significant differences between MNA-SF categories with energy and protein intake were identified. **Conclusion:** Energy and protein intake were inadequate among older adults in this present study. Malnutrition risk was prevalent in this population, with older age groups being significantly at risk. Strategies to address nutritional issues among Malaysian community older adults are warranted for optimum nutritional and health status.

INTRODUCTION

Ageing is a global phenomenon that has various implications on health, with significant impacts on medical care and health policy (Noto, 2023). According to the World Health Organization (WHO), one in six people worldwide will be aged 60 years or older (WHO, 2025). Furthermore, the population aged 60 years and above is expected to increase from 1 billion in 2020 to 1.4 billion in 2030. In addition, by 2050, this population is expected to double (2.1 billion). The number of people aged 80 years and older is expected to triple between 2020 and 2050, reaching 426 million. In Malaysia, the proportion of older adults aged 60 years and above is also projected to accelerate significantly to 15.3% in 2030, to approximately 5.8 million (Abdullah et al., 2024). Malaysia is expected to become an ageing nation by 2030, with significant implications for health services and support systems (DOSM, 2023).

Ageing is a biological process defined as a progressive loss of function accompanied by decreasing fertility, increasing death, and possible diseases (Kirkwood & Austad, 2000; Viña et al., 2007). As the population ages, maintaining adequate nutritional status becomes a crucial component of healthy ageing. Energy and protein intake are particularly important among older adults, as they play vital roles in maintaining muscle mass, immune function, and overall health (Bauer et al., 2013; Volkert et al., 2019). Insufficient intake of these nutrients may lead to undernutrition, sarcopenia, frailty, and increased morbidity and mortality (Norman et al., 2021).

Older adults living in the community often face multiple challenges that can compromise their nutritional intake. Factors such as physiological changes, reduced appetite, poor dentition, limited income, social isolation, and chronic illnesses can influence dietary patterns and nutrient adequacy (Park & Kang, 2024). Hence, understanding the energy and protein intake among this group is essential for developing appropriate public health

* *Corresponding author.*
E-mail address: aliza@iium.edu.my

strategies and community interventions to promote healthy ageing.

Several studies in Malaysia have reported alarming concerns regarding inadequate dietary intake and malnutrition risk among older adults. Suzana et al. (2013), reported that 43.1% male and 42.1% female elderly participants were at risk of malnutrition according to the Mini Nutritional Assessment-Short Form (MNA-SF). Similarly, Abdul Mutalib et al. (2023) found that 11% of elderly participants were malnourished and 48% were at risk of malnutrition, despite having higher energy consumption compared to the Recommended Nutrient Intake (RNI). In another study, Ahmad et al. (2021) found that sociodemographic characteristics such as rural residence and low education level were associated with malnutrition risk among older adults, with the cumulative prevalence of malnutrition and risk of malnutrition of 30.8%.

To assess the malnutrition risk in older adults, validated screening tools such as the Mini Nutritional Assessment–Short Form (MNA-SF) are required. MNA-SF is a short, valid screening tool that has been developed to identify malnutrition in a community setting (Kaiser et al., 2009). The MNA-SF consists of 6 items that have been developed from 18 items of the full MNA, but still retain the accuracy and validity (Rubenstein et al., 2001). It evaluates several key parameters, including food intake, weight loss, mobility, psychological stress, neuropsychological problems, and body mass index (BMI) or calf circumference (Guigoz, 2006). This tool is particularly suitable for community settings as it is simple, quick to administer, and has been validated in diverse populations, including Malaysian older adults (Shahar & Siti Saifa, 2007).

As Malaysia faces the demographic shift, assessing the dietary intakes, specifically energy and protein, and malnutrition risk using MNA-SF is needed for a comprehensive understanding of the nutritional well-being of older adults. Combining these assessments helps identify those at risk of malnutrition and informs targeted interventions to improve dietary practices and overall health outcomes. Despite existing national data, there is limited evidence that specifically explores energy and protein intake with malnutrition risk among urban older adults within a local context in Pahang. Therefore, this study aimed to determine the energy and protein intake and evaluate its association with malnutrition risk among community-living older adults in Kuantan, Pahang, using the MNA-SF tool.

MATERIALS AND METHODS

Study Design and Population

A cross-sectional study was conducted among urban community-living older adults in Kuantan, Pahang. The participants who were 60 years old and above, living in Kuantan, Pahang, and willing to participate in this study were selected.

Sampling Method

According to the Department of Statistics Malaysia (DOSM, 2020), Kuantan has a population of 128,247 with 5.3% of older adults. Hence, the estimation of sample size was calculated by using a formula (Naing et al., 2006) where the standard value of confidence level 95% was set as 1.96, absolute precision was set as 5% and the estimated prevalence of malnutrition among community living older adults in Kuantan was 5%. Therefore, after calculating, a total of seventy-three (n=73) respondents were required as the sample size in this study.

A convenient sampling method was conducted for the recruitment of the participants. Thus, participants who fulfilled the inclusion criteria were provided with an informed consent form to fill in for their voluntary participation. The respondents were informed that their identities would remain anonymous and that the data collected would be used for research purposes only.

Data Collection

The participants were informed about this study. Hence, the participants who voluntarily participated in this study were interviewed on their demographic data, diet history, and MNA-SF questions.

Anthropometric measurements

The weight and height of respondents were measured using a weighing scale (SECA 803, Hamburg, Germany) and a portable stadiometer (SECA 213, Hamburg, Germany), respectively. Measurements of weight (kilogram) and height (metre) were needed to calculate body mass index (BMI). Furthermore, weight changes of the respondents were also asked to identify the amount of weight loss, if any. This information was collected to fill in the MNA-SF questionnaire.

Dietary intake assessment

The dietary intake assessment was collected face-to-face with the elder respondents. The questions that were asked

were about mealtimes, place of eating, food intake, types of food, amount or quantity of food taken, and the most frequent food intake. The diet history method was utilized as it provides both information, dietary intake and food habits of the individuals, and it shows a better understanding of undernutrition and overnutrition (Shahar et al., 2000). The diet history was analysed by using Nutritionist Pro Software version 2006 to obtain the total energy and protein intake.

Next, open-ended questions were asked to explore participants' protein intake preference, including whether they commonly consumed animal-based or plant-based protein sources, and the frequency of intake for specific foods such as beef or lamb, chicken, fish, and legumes. The common sources of protein were then identified based on the reported frequency of consumption per week.

MNA-SF questionnaire

The MNA-SF was administered by a researcher. It is a validated tool that consists of 6 questions and takes only 5 minutes to complete (Kaiser et al., 2009) and has been validated in the Malaysian older adults population (Suzana & Siti Saifa, 2007). The components of MNA-SF are as follows:

A. Has food intake declined over the past 3 months due to loss of appetite, digestive problems, or chewing or swallowing difficulties?
B. Weight loss during the last 3 months
C. Mobility
D. Has suffered psychological stress or acute disease in the past 3 months?
E. Neuropsychological problems
F. Body Mass Index (BMI) OR
F2. Calf Circumference

MNA-SF has a maximum score of 14. Participants' malnutrition risk has been categorized into three groups: well-nourished (score 12-14 points), at-risk of malnutrition (8-11 points), and malnourished (0-7 points) (Kaiser et al., 2009). Patient has satisfactory nutritional status if the score is ≥12 and 14.

Data Analysis

Data was analysed using the Statistical Package for Social Sciences (SPSS) version 12.0. Next, descriptive analysis was used to obtain the mean and standard deviation (SD) for age, energy, and protein intake. Moreover, the comparison of energy and protein intake with recommended nutrient intake (RNI) was analysed by using a sample T-test. The association between energy and protein intake with MNA-SF score categories (normal nutritional status, at risk of malnutrition, and malnourished) was analysed using One-Way ANOVA. The significance level was set at $p<0.05$.

Ethics Application

Ethics approval was obtained from the International Islamic University Malaysia (IIUM) Research Ethics Committee (IREC) (IREC NO: 618).

RESULTS

Demographic Data

A total of seventy-three (n=73) respondents from Kuantan were recruited for this research. From the total respondents, there were n=37 (50.7%) male and n=36 (49.3%) female older adults. The age of the respondents ranges from 60 - 84 years old. All respondents were Malay. Table 1 shows the age and BMI of the respondents. The mean age was 65.88±5.67. The mean BMI was 26.023±4.21, and the BMI ranged from 19.5 to 36.8 kg/m^2.

Table 1: Demographic data of the respondents (n=73)

Variable	Mean±SD	Range
Age in years	65.88±5.67	60-84
BMI (kg/m^2)	26.023±4.21	19.5-36.80

Dietary Intake

Table 2 presents the mean energy and protein intake per day compared with the Recommended Nutrient Intakes (RNI) for Malaysian elderly. The mean±SD energy intake of the respondents was 1287.50±404.11, which was significantly lower than the recommended energy requirement by RNI (1896.58±115.79). Besides, the protein intake, 47.40±14.28, was also significantly lower than the recommended protein requirement by RNI (55.10±4.03). The one-sample t-test showed that the results demonstrated significant differences, as the *p*-value was <0.001 (*p*-value<0.05) for both energy and protein intake compared with RNI.

Table 2: Mean±SD energy and protein intake compared with RNI

Variable	Intake Mean±SD	RNI Mean±SD	*p*-value*
Energy (kcal/day)	1287.50±404.11	1896.58±115.79	<0.001
Protein (g/day)	47.40±14.28	55.10±4.03	<0.001

*One Sample t-test

Dietary Practices

Table 3 shows the frequencies and percentages of respondents' dietary practices related to place of eating, preferred type of protein, and preferred protein sources. 84.9% (n=62) of the respondents eat at home, while 15.1% (n=11) of the respondents practice eating out. 98.6% of respondents preferred animal-based protein sources, whilst only 1.4% preferred legumes and soy products. The most preferred protein source by respondents was fish (80.8%). While other preferred protein sources by the respondents were: chicken (15.1%), beef/lamb (2.7%), and 1.4% preferred legumes and soy products.

Table 3: The dietary practices of respondents (n=73)

Dietary practices	Frequency (n)	Percentages (%)
Place of eating		
- Home	62	84.9
- Outside	11	15.1
Preferred type of protein		
- Animal	72	98.6
- Legumes, soy product	1	1.4
Preferred protein sources		
- Beef/lamb	2	2.7
- Chicken	11	15.1
- Fish	59	80.8
- Legumes and soy products	1	1.4

The Association between Age, Energy, and Protein Intake with MNA-SF score

Table 4 shows the association between age, energy, and protein intake with the MNA-SF score. According to the MNA-SF score, 11% were malnourished, 52% at risk of malnutrition (n=38), and 37% were in the normal nutritional status category, respectively. From the MNA-SF score, the results indicated malnutrition risk status of the respondents was associated with the respondents with advancing age, who were identified as malnourished and had poor nutrition. The respondents with a mean age of 71.63±4.27 were in the malnourished category, 66.47±6.11 were at risk of malnutrition, whilst 63.33±3.70 were in the normal nutritional status category, respectively. There was a statistically significant difference between age and the MNA-SF screening score category as determined by one-way ANOVA with p-value<0.001. A Tukey post hoc test revealed that the age in the MNA-SF score was significantly different among normal nutritional status and at-risk-of-malnutrition respondents (p=0.047), normal nutritional status and malnourished respondents (p<0.001), and at-risk-of-malnutrition and malnourished respondents (p=0.033).

Table 4: The association between age, energy, and protein intake with the MNA-SF score

Variables	Malnourished (n=8)	At risk of malnutrition (n=38)	Normal nutrition status (n=27)	Levene's test *p*-value*
Age in years mean±SD	71.63±4.27	66.47±6.11	63.33±3.70	<0.001
Energy (kcal) mean±SD	1064.52±387.21	1305.95±408.46	1327.60±396.18	0.132
Protein (g) mean±SD	43.33±15.72	47.41±14.31	48.59±14.16	0.229

*One-Way ANOVA

The mean energy intake of the respondents in the malnourished category was 1064.52±387.21, and at risk of malnutrition was 1305.95±408.46, which were lower than the mean energy intake of respondents in the normal nutrition status category (1327.60±396.18). Nevertheless, there was no significant difference between energy intake

with MNA-SF score categories (malnourished, at risk of malnutrition, and normal nutrition status) (p=0.132). Besides, no significant difference was detected between protein intake with MNA-SF score (p=0.229).

DISCUSSION

This study assessed energy and protein intake among community living older adults in Kuantan, Pahang and its association with malnutrition risk was explored.

Energy and Protein Intake

This study revealed that the energy and protein intake were lower than the recommendation. The findings are in line with a research among Malaysian community living older adults in agricultural settlements (Zainudin, Hamirudin, Sidek, & A. Rahman, 2020). Another study among older adults in rural regions of Peninsular Malaysia also reported inadequate energy intake, although a notable higher intake among male (1412 ± 461 kcal/day) than female (1201 ± 392 kcal/d) was identified (Shahar et al., 2007). A more recent study by Ja'afar et al. (2024) reported 34.9% Malaysian community living elderly had lower energy intake than recommendation, with the use of different dietary methodology using food frequency questionnaire in comparison to diet history method in our study.

Adequate energy and nutrients intake is essential to ensure optimum nutritional status to prevent health deterioration, which could lead to poor health outcomes, hospitalization and mortality (Norman, Haß, & Pirlich, 2021). The multiple contributors to insufficient intake are physiological changes in which anorexia of aging predominates, diminished appetite (Picca, Calvani, Coelho-Júnior, Landi, & Marzetti, 2022; Tsutsumimoto et al., 2020). The anorexia of ageing is synonymous with older adults, in which declining dietary intake is a notable phenomenon in this vulnerable population (Picca et al., 2022).

Poor appetite has been documented as a predictor of low energy intake among Malay older adults residing in an urban region (Mohamad, Suzana, Noor Ibrahim, & Norshafarina, 2010), which shared similar demographics with our studied population. Prolonged energy deficit could further lead to protein-energy malnutrition, which is characterized by deterioration of both protein and energy intake, a significant issue among older adults. A pioneer study in Malaysia demonstrated that both protein and energy intake of community living older adults are inadequate, with significant deterioration across older age groups (Suriah et al., 1996). This condition not only deteriorates older adults' nutritional status but also associated with decreased muscle mass and functional impairment (Agarwal, Miller, Yaxley, & Isenring, 2013). The adverse effects include sarcopenia; a low muscle mass with poor muscle strength with a greater susceptibility to fall among older adults (Zhong et al., 2022).

Inadequacy of protein intake in this present study is consistent with findings in other study among Malaysian community living older adults (Zainudin et al., 2020). Another Malaysian study assessing nutrient deficiencies in this vulnerable population indicating an alarming rate of 85% inadequate protein intake (Ja'afar et al., 2024). Nevertheless, contrasting findings were reported by Mohd Fakhruddin, Shahar, Aziz, Yahya, and Rajikan (2016); in which most elder subjects have adequate intake particularly among female elderly. Protein is a vital nutrient for muscle protein synthesis, with a significant role in functional status and independent activities of daily living (Alamilla, Paulussen, Askow, & Burd, 2021; Nunes, Currier, Lim, & Phillips, 2021). Suboptimal protein intake is often accompanied with other nutrient deficiencies that could lead to frailty, poor cognition and impaired quality of life (Del Carmen Alvarez-Nuncio & Ziegler, 2024).

Dietary Practices of Older Adults

Interestingly, our research revealed that older adults have a strong preference for fish and a low preference for plant-based sources. Findings in a study among the majority Malay older adults in Mukim Plentong, Johor Bahru, also showed that fish was the highest protein consumed, including eggs and anchovies (Fadzwi, Sulaiman, Ibrahim, & Appannah, 2025). This can be attributed to the sociocultural acceptance of fish among the Malay community in Malaysia and its texture, which requires less chewing than meat and poultry. Regular fish consumption provides essential amino acids and is are good source of omega-3 fatty acids; beneficial for preserving muscle mass and metabolic health (Mendivil, 2021). Moreover, legumes and soy products are less synonymous among Malays as habitual protein sources. Notably, research on the type of protein sources consumed among older adults is scarce in Malaysia, which warrants further investigation. Results

from the Malaysian Adults Nutrition Survey (MANS) among the population aged 18-59 years old reported that most respondents residing in urban regions consumed chicken and eggs, whilst marine fish sources were recorded as higher among urban (51%) than rural (34%) counterparts (Norimah et al., 2008).

Eating at home is the most frequent practice of our studied population than eating out, in line with previous research (Zainudin, Hamirudin, A. Rahman, & Sidek, 2019). The cost of eating out is commonly higher than self-prepared food at home. Nevertheless, a balanced diet is an utmost importance in daily intake. A systematic review highlighted that home-delivered meals services can improve energy and protein intake in older adults, which could prevent further nutrition and health deterioration (Walton, Rosario, Pettingill, Cassimatis, & Charlton, 2019). This approach could be an alternative to improve dietary intake among Malaysian older adults.

The Association between Age, Energy, and Protein Intake with MNA-SF score

This present study demonstrated that 11% were malnourished, whilst 52% were at risk of malnutrition. Meanwhile, the most recent Malaysian nationwide study in older adults reported a lower prevalence using a similar MNA-SF screening tool with 7.3% malnutrition and 23.5% at risk of malnutrition, respectively (Ahmad et al., 2021). In agricultural settlements, Suzana, Boon, Chan, and Normah (2013) revealed a 42.5% of older adults were at risk malnutrition in Selangor; whilst lower prevalence of 25.7% was identified by Zainudin et al. (2019) in Pahang. Meanwhile, Sheikh Hishamuddin et al. (2023) reported 30% were malnourished and at-risk among older adults attending health clinics. A pilot study in similar population origin documented 64% malnutrition risk (Muhamad, Hamirudin, Zainudin, Sidek, & A. Rahman, 2019).

Our study demonstrated that advancing age is significantly associated with malnutrition risk. The finding is parallel with other Malaysian studies in agricultural settlements (Zainudin et al., 2019) and another study in a health clinic setting (Sheikh Hishamuddin et al., 2023). Previous research indicated that older adults with financial limitations is associated with malnutrition risk (Zainudin et al., 2019). Nevertheless, energy and protein intake showed no significant difference with MNA-SF categories in the present study. The findings highlight the multifactorial influences and determinants of malnutrition risk beyond dietary intake, with other well-documented potential contributors related to physiological, psychological, functional, and social support (Tomasiewicz, Polański, & Tański, 2024).

A nationwide survey demonstrated that food insecurity is a predictor of malnutrition risk among older adults, which is related to food accessibility and affordability (Ahmad et al., 2021). Food insecurity is also associated with malnutrition risk, level of education, living in a rural region, and income (Salleh et al., 2020) in which food cost is prioritized over nutrient-dense food in dietary intake. Several other factors may also be simultaneously contributing to malnutrition, such as lack of appetite, diminished taste and smell, presence of acute or chronic diseases, and poor dietary intake (Norman et al., 2021). Although the mean BMI of this studied population was in the recommended range, within 24- 27 kg/m^2 for Malaysian elderly (Ministry of Health, 2023), this clearly indicates that BMI is not a sole indicator of nutritional status in older adults. To address malnutrition risk on time, nutrition screening within a community setting is required, and health clinic settings have been identified as ideal (Sheikh Hishamuddin et al., 2023). Nutrition intervention using a specifically tailored nutrition resource kit for the malnourished and at-risk older adults was well-accepted by the older adults themselves (Noor Azam et al., 2022).

Limitations and Strengths

The limitations of this study are its preliminary characteristics, which limit the generalizability to a larger population. The cross-sectional study design also limits the causality and precludes inference of dietary intake and malnutrition risk. The identified strength of this study is the use of the diet history method to assess usual dietary intake in comparison to other methods, which might not be suitable for older adults, although a recall bias might be present. Moreover, the MNA-SF tool is a validated tool for older adults with high sensitivity and specificity to identify those who are at malnutrition risk was an instrument utilized in this study. It is worth noting that this study provides insights into the energy and protein intake of older adults' dwellers in Kuantan, an underrepresented urban region in Malaysia's aging research landscape.

CONCLUSION

This study demonstrates that the energy and protein intake of older adults in Kuantan were significantly lower than recommendations. Most respondents practiced eating at home and indicated a preference for animal-based protein sources, particularly fish. There was a statistically significant difference between age and the MNA-SF categories, but no significant difference between energy and protein intake. For future research, a larger sample size and inclusion of multiethnicity are recommended. In addition, effective nutrition education strategies are essential to ensure optimum nutritional

status among community-living older adults along with timely nutrition screening and intervention.

ACKNOWLEDGEMENT

We sincerely thank all respondents for their voluntary participation in this this research.

REFERENCES

Abdul Mutalib, M., Mohd Noor, N., Shamsuddin, A. S., Abdul Kadir, N. A. A., & Adam, S. H. (2023). Assessment Nutritional Status of Elderly People at Old Folk Home in Klang Valley. *Journal of Management & Science*, *21*(1), 12. https://doi.org/10.57002/jms.v21i1.364

Abdullah, J. M., Ismail, A., & Yusoff, M. S. B. (2024). Healthy Ageing in Malaysia by 2030: Needs, Challenges and Future Directions. In *Malaysian Journal of Medical Sciences* (Vol. 31, Issue 4, pp. 1–13). Penerbit Universiti Sains Malaysia. https://doi.org/10.21315/mjms2024.31.4.1

Agarwal, E., Miller, M., Yaxley, A., & Isenring, E. (2013). Malnutrition in the elderly: A narrative review. *Maturitas, 76*(4), 296-302. doi:http://dx.doi.org/10.1016/j.maturitas.2013.07.013

Ahmad, M. H., Salleh, R., Siew Man, C., Pardi, M., Che Abdul Rahim, N., Shahril, N., . . . Ahmad, N. A. (2021). Malnutrition among the Elderly in Malaysia and Its Associated Factors: Findings from the National Health and Morbidity Survey 2018. *Journal of Nutrition and Metabolism, 2021*. doi:10.1155/2021/6639935

Alamilla, R. A., Paulussen, K. J. M., Askow, A. T., & Burd, N. A. (2021). Dietary approaches to maintaining muscle mass. In *Sarcopenia: Molecular Mechanism and Treatment Strategies* (pp. 81-107): Elsevier.

Bauer, J., Biolo, G., Cederholm, T., Cesari, M., Cruz-Jentoft, A. J., Morley, J. E., Phillips, S., Sieber, C., Stehle, P., Teta, D., Visvanathan, R., Volpi, E., & Boirie, Y. (2013). Evidence-based recommendations for optimal dietary protein intake in older people: A position paper from the prot-age study group. *Journal of the American Medical Directors Association*, *14*(8), 542–559. https://doi.org/10.1016/j.jamda.2013.05.021

Del Carmen Alvarez-Nuncio, M., & Ziegler, T. R. (2024). Micronutrient status and protein-energy malnutrition in free-living older adults: A current perspective. *Current Opinion in Gastroenterology, 40*(2), 99-105. doi:10.1097/MOG.0000000000001000

Department of Statistics Malaysia. (2023). *Current Population Estimates, Malaysia, 2023* [PDF]. https://www.dosm.gov.my/site/downloadrelease?admin_view=&id=current-population-estimates-malaysia----2023&lang=English

Department of Statistics Malaysia. (2020). *Kawasanku dashboard: P.083 Kuantan parliamentary constituency, Pahang* [Data dashboard]. https://open.dosm.gov.my/dashboard/kawasanku/Pahang/parlimen/P.083%20Kuantan

Fadzwi, F. M., Sulaiman, N., Ibrahim, Z., & Appannah, G. (2025). Adherence of Malaysian Older Persons to the ''Variety''Characteristic based on the Malaysian Dietary Guidelines. *International Journal, 2*(2), 54-62.

Guigoz, Y. (2006). The Mini Nutritional Assessment (MNA(registered trademark)) review of the literature - What does it tell us? *Journal of Nutrition, Health and Aging*, *10*(6), 466–485. https://doi.org/10.1016/S0899-9007(98)00171-3

Ja'afar, M. H., Mat Nasir, N., Md Isa, Z., Ismail, R., Mohd Tamil, A., Ismail, N. H., . . . Yusof, K. H. (2024). Dietary nutrient intake study among older adults: baseline Malaysian pure study. *BMC Geriatrics, 24*(1), 441. doi:10.1186/s12877-024-05042-w

Kaiser, M. J., Bauer, J. M., Ramsch, C., Uter, W., Guigoz, Y., Cederholm, T., Thomas, D. R., Anthony, P., Charlton, K. E., Maggio, M., Tsai, A. C., Grathwohl, D., Vellas, B., & Sieber, C. C. (2009). Validation of the Mini Nutritional Assessment short-form (MNA®-SF): A practical tool for identification of nutritional status. *The Journal of Nutrition, Health and Aging*, *13*(9), 782–788. https://doi.org/10.1007/s12603-009-0214-7

Kirkwood, B., & Austad, S. (2000). Why do we age? In *Nature* (Vol. 408). https://doi.org/10.1038/35041682

Mendivil, C. O. (2021). Fish Consumption: A Review of Its Effects on Metabolic and Hormonal Health. *Nutrition and Metabolic Insights, 14*. doi:10.1177/11786388211022378

Ministry of Health. (2023). *Malaysian Dietary Guideline for Older Persons*. In National Coordinating Committee for Food and Nutrition (NCCFN) (Ed.). Retrieved from https://hq.moh.gov.my/nutrition/wp-content/uploads/2023/12/Buku-MDGOP-1.pdf

Mohamad, H. A., Suzana, S., Noor Ibrahim, M. S., & Norshafarina, S. (2010). Relationship between appetite, food intake and body composition among elderly Malays from an urban residential area in Kuala Lumpur, Malaysia. *Malaysian Journal of Nutrition, 16*(3), 339-348. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-79851471273&partnerID=40&md5=f2b3fb2de61248f5a43b3386ac959f14

Mohd Fakhruddin, N. N. I. N., Shahar, S., Aziz, N. A. A., Yahya, H. M., & Rajikan, R. (2016). Which aging group prone to have inadequate nutrient intake?: TUA Study (Kumpulan penuaan yang mana lebih cenderung terhadap pengambilan nutrien yang tidak mencukupi?: Kajian TUA). *Sains Malaysiana, 45*(9), 1381-1391. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84992058714&partnerID=40&md5=4c0feb3fa7398d6071ea0343e49df78b

Muhamad, A. R., Hamirudin, A. H., Zainudin, N., Sidek, S., & A. Rahman, N. A. (2019). Nutritional Risk According To Mini Nutritional Assessment-Short Form Among Community Dwelling Elderly In Kuantan, Pahang: A Pilot Study. *International Journal Of Allied Health Sciences, 3*(2), 658–667.

Naing, L., Naing, L., Winn, T., & Rusli, B. N. (2006). Pratical Issues in Calculating the Sample Size for Prevalence Studies Practical Issues in Calculating the Sample Size for Prevalence Studies. In *Archives of Orofacial Sciences* (Vol. 1). https://www.researchgate.net/publication/45459002

Noor Azam, I. N., Hamirudin, A. H., Harith, S., Md Aris, M. A., Abd Aziz, K. H., & A Rashid, N. S. (2022). Development, validation and acceptability of a newly developed nutrition resource kit for at-risk and malnourished elderly in health clinics setting. *Malaysian Journal of Medicine and Health Sciences, 18*(5), 93-103.

Norimah, A. K., Jr., Safiah, M., Jamal, K., Haslinda, S., Zuhaida, H., Rohida, S., . . . Azmi, M. Y. (2008). Food Consumption Patterns: Findings from the Malaysian Adult Nutrition Survey (MANS). *Malays J Nutr, 14*(1), 25-39.

Norman, K., Haß, U., & Pirlich, M. (2021). Malnutrition in older adults-recent advances and remaining challenges. *Nutrients, 13*(8). doi:10.3390/nu13082764

Noto, S. (2023). Perspectives on Aging and Quality of Life. In *Healthcare (Switzerland)* (Vol. 11, Issue 15). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/healthcare11152131

Nunes, E. A., Currier, B. S., Lim, C., & Phillips, S. M. (2021). Nutrient-dense protein as a primary dietary strategy in healthy ageing: Please sir, may we have more? *Proceedings of the Nutrition Society, 80*(2), 264-277. doi:10.1017/S0029665120007892

Park, J. H., & Kang, S. W. (2024). Nutritional Risk, Depression, and Physical Function in Older People Living Alone. *Healthcare (Switzerland), 12*(2). https://doi.org/10.3390/healthcare12020164

Picca, A., Calvani, R., Coelho-Júnior, H. J., Landi, F., & Marzetti, E. (2022). Anorexia of Aging: Metabolic Changes and Biomarker Discovery. *Clinical interventions in aging, 17*, 1761-1767. doi:10.2147/cia.S325008

Rubenstein, L. Z., Harker, J. O., Salvà, A., Guigoz, Y., & Vellas, B. (2001). *Screening for Undernutrition in Geriatric Practice : 56*(6), 366–372.

Salleh, R., Man, C. S., Ahmad, M. H., Palaniveloo, L., Zulkafly, N., Ab Halim, S. A., . . . Ibrahim Wong, N. (2020). Factors contributing to food insecurity among older persons in Malaysia: Findings from the National Health and Morbidity Survey (NHMS) 2018. *Geriatr Gerontol Int, 20 Suppl 2*, 73-78. doi:10.1111/ggi.14006

Shahar, S., Earland, J., & Rahman, S. A. (2000). Food intakes and habits of rural elderly Malays. *Asia Pacific Journal of Clinical Nutrition*, *9*(2), 122–129. https://doi.org/10.1046/j.1440-6047.2000.00147.x

Shahar, S., Ibrahim, Z., Fatah, A. R. A., Rahman, S. A., Yusoff, N. A. M., Arshad, F., . . . Adznam, S. N. (2007). A multidimensional assessment of nutritional and health status of rural elderly Malays. *Asia Pacific Journal of Clinical Nutrition, 16*(2), 346-353. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-34248561108&partnerID=40&md5=71a1e360f2d367141485e9ec07f3cc6c

Sheikh Hishamuddin, S. N. A., Hamirudin, A. H., Harith, S., Md Aris, M. A., Abd Aziz, K. H., & A Rashid, N. S. (2023). Development and Feasibility Assessment of a Nutrition Screening Guideline in Older Patients for Use by

Healthcare Staff in Health Clinic Setting: An Action Research Study. *Malaysian Journal of Medicine and Health Sciences, 19*(1), 46-56. doi:10.47836/mjmhs.19.1.8

Suriah, A., Zainorni, M., Shafawi, S., Mimie Suraya, S., Zarina, N., Wan Zainuddin, W., & Zalifah, M. (1996). Nutrient intake among elderly in southern Peninsular Malaysia. *Malaysian Journal of Nutrition, 2*(1), 11-19.

Suzana, S., Boon, P., Chan, P., & Normah, C. (2013). Malnutrition risk and its association with appetite, functional and psychosocial status among elderly Malays in an agricultural settlement. *Malaysian Journal of Nutrition, 19*(1), 65-75.

Suzana, S., & Siti Saifa, H. (2007). Validation of Nutritional Screening Tools Against Anthropometric and Functional Assessments Among Elderly People in Selangor. *Mal J Nutr, 13*(1), 29-44.

Tomasiewicz, A., Polański, J., & Tański, W. (2024). Advancing the Understanding of Malnutrition in the Elderly Population: Current Insights and Future Directions. *Nutrients, 16*(15). doi:10.3390/nu16152502

Tsutsumimoto, K., Doi, T., Nakakubo, S., Kim, M., Kurita, S., Ishii, H., & Shimada, H. (2020). Association between anorexia of ageing and sarcopenia among Japanese older adults. *Journal of Cachexia, Sarcopenia and Muscle, 11*(5), 1250-1257. doi:10.1002/jcsm.12571

Viña, J., Borrás, C., & Miquel, J. (2007). Theories of ageing. *IUBMB Life*, *59*(4–5), 249–254. https://doi.org/10.1080/15216540601178067

Volkert, D., Beck, A. M., Cederholm, T., Cruz-Jentoft, A., Goisser, S., Hooper, L., Kiesswetter, E., Maggio, M., Raynaud-Simon, A., Sieber, C. C., Sobotka, L., van Asselt, D., Wirth, R., & Bischoff, S. C. (2019). ESPEN guideline on clinical nutrition and hydration in geriatrics. *Clinical Nutrition*, *38*(1), 10–47. https://doi.org/10.1016/j.clnu.2018.05.024

Walton, K., Rosario, V. d., Pettingill, H. S., Cassimatis, E., & Charlton, K. (2019). The Impact of Home-delivered Meal Services on the Nutritional Intake of Community Living Older Adults: A Systematic Literature Review. *Journal of Human Nutrition and Dietetics, 33*(1), 38-47. doi:10.1111/jhn.12690

World Health Organization. (2025). *Ageing and health* [Fact sheet]. https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

Zainudin, N., Hamirudin, A. H., A. Rahman, N. A., & Sidek, S. (2019). Malnutrition risk and perception on dietary practices among elderly living in agricultural settlements: A mixed-methods research. *Nutrition and Food Science, 49*(4), 617-627. doi:10.1108/NFS-07-2018-0218

Zainudin, N., Hamirudin, A. H., Sidek, S., & A. Rahman, N. A. (2020). Dietary intake is compromised among elderly living in agricultural settlements. *Nutrition and Food Science, 50*(2), 314-323. doi:10.1108/NFS-01-2019-0028

Zhong, J., Xie, W., Wang, X., Dong, X., Mo, Y., Liu, D., . . . Wang, X. (2022). The Prevalence of Sarcopenia Among Hunan Province Community-Dwelling Adults Aged 60 Years and Older and Its Relationship With Lifestyle: Diagnostic Criteria From the Asian Working Group for Sarcopenia 2019 Update. *Medicina, 58*(11), 1562. doi:10.3390/medicina58111562

ORIGINAL ARTICLE

Anthropometric Profile, Sports Nutrition Knowledge and Its Association Among Karate Athletes in Kuantan, Pahang

Muhammad Amirul Aiman Muhammad Yuswadi [1], Norashikin Mustafa[1,2*]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Kuantan, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Kuantan, Pahang, Malaysia

Keywords:
karate athletes; somatotype; sports nutrition knowledge; body fat; anthropometry

ABSTRACT

Background: Karate is a growing combat sport in Malaysia with increasing international success. Anthropometric profiles such as somatotype, body mass index (BMI), and body fat percentage influence combat performance, while sports nutrition knowledge (SNK) may further affect these factors. Therefore, this study aimed to assess the anthropometric profiles, SNK level, and its associations among karate athletes in Kuantan, Pahang. **Methods:** A total of 26 male and female karate athletes aged 14–21 years from Majlis Sukan Pahang participated in the study. BMI was calculated from height and weight, body fat percentage was estimated using two-site skinfold measurements and the Slaughter et al. equation, and somatotype was determined using the Carter and Heath method. SNK was assessed using the Nutrition Knowledge for Young Athletes (NUKYA) questionnaire. Spearman's rank-order correlation was employed to analyze associations between anthropometric profile and SNK. **Results:** Male athletes had a mean BMI of 21.3 ± 2.95 kg/m², body fat percentage of 21%, and predominantly mesomorphic somatotype. Female athletes recorded a mean BMI of 21.6 ± 0.34 kg/m², body fat percentage of 24.2 ± 3.43%, and predominantly endomorphic characteristics. Overall, athletes demonstrated low SNK, with only 11.5% scoring above 60%. No significant association was found between somatotype and SNK in either gender. However, a significant negative correlation was observed between BMI and SNK among females ($r = –0.845$, $p < 0.0001$), while a significant positive correlation was found between body fat percentage and SNK among males ($r = 0.427$, $p = 0.007$). **Conclusion:** Malaysian karate athletes in this study exhibited generally low levels of sports nutrition knowledge. While SNK showed associations with BMI (among females) and body fat percentage (among males), no significant relationship was observed with somatotype across gender. These findings suggest that nutrition knowledge may not strongly influence inherent body type but underscore the importance of targeted nutrition education programs to improve dietary practices and athletic performance. Further research involving larger karate athlete populations is recommended to validate and expand upon these findings.

INTRODUCTION

Karate is one of the most popular martial arts or combat sports that have been practiced all around the world (Koropanovski et al., 2011). Karate sports competition can be further divided into two distinguished disciplines, which are kata and kumite. Kata is a set of predetermined movements involving offensive and defensive action in standard order, against a non-existent opponent; meanwhile, kumite represents combat activities of karate between two karatekas under certain rules and a scoring system (Slankamenac et al., 2021).

Karate is one of Malaysia's most prominent combat sports, alongside silat, taekwondo, and boxing. Governed by the Malaysia Karate Federation (MAKAF), the sport has produced several international medalists at the 19th Asian Games in Hangzhou (2022). This achievement reflects Malaysia's growing presence in the regional karate scene and underscores the importance of understanding factors influencing athlete performance, including anthropometric profiles and sports nutrition knowledge (Malaysia Olympic Council, 2023)

Athletes need to possess specific physical attributes to perform optimally in their respective sports, including karate, which involves direct physical contact between opponents. To adapt to the demands of specific disciplines, athletes are required to exhibit somatic characteristics that align with the level and intensity of their sport (Lijewski et al., 2021). Previous studies have shown that most karate athletes, regardless of gender, exhibit an endomorphic–mesomorphic somatotype, characterized by higher muscularity and lower body fat (Slankamenac et al., 2021). In addition, anthropometric profiles play a crucial role in determining training methods and intensities that support

* *Corresponding author.*
E-mail address: norashikinmustafa@iium.edu.my

the physical fitness and strength requirements unique to each sport (Lijewski et al., 2021). Competing at an elite level requires athletes to maintain optimal body composition and physiological characteristics throughout the competitive season (Samanes et al., 2022). Given the strength and physical demands inherent in karate, athletes must develop and sustain specific anthropometric and somatotype profiles to enhance performance. However, most existing research emphasizing these parameters focuses on team sports, leaving a notable gap in understanding the anthropometric and somatotype requirements of combat sport athletes such as karate.

Recent studies have highlighted the importance of sports nutrition knowledge among athletes (Alahmadi & Albassam, 2023; Thapa et al., 2023). Adequate knowledge in sports nutrition is essential as it influences dietary intake and eating habits, thereby enhancing athletic performance (Alahmadi & Albassam, 2023). Combat sports, including karate, employ a weight division system to ensure fair competition between opponents. However, athletes often resort to acute weight loss (AWL) strategies, such as rapid dehydration, to qualify for lower weight categories (Reale, 2018). Therefore, it is crucial for karate athletes to possess adequate knowledge of macronutrient intake to maintain a balanced diet, achieve optimal body composition, and sustain peak performance (Reale, 2018). Furthermore, inadequate nutrition knowledge may lead to inappropriate use of supplements and dietary practices that do not align with athletes' performance goals (Alahmadi & Albassam, 2023). Nevertheless, research on sports nutrition knowledge among Malaysian karate athletes remains scarce.

Assessing athletes' anthropometric profiles is essential to determine the body morphology most suited to their sport. Previous studies have shown that appropriate anthropometric and somatotype characteristics enhance performance, reduce injury risk, and improve technical skills (Bekele et al., 2022; Čaušević et al., 2023; Kastrati et al., 2022; Slankamenac et al., 2021). However, limited research has explored the relationship between anthropometric profiles and sports nutrition knowledge, particularly among Malaysian athletes. Addressing this gap, the present study aims to identify the anthropometric profile, sports nutrition knowledge, and their association among karate athletes in Kuantan, Pahang, to better understand how nutrition knowledge may influence body composition.

MATERIALS AND METHODS

Research Design

A cross-sectional study was conducted to examine the association between anthropometric profiles and sports nutrition knowledge among karate athletes in Kuantan, Pahang. This design was selected to describe the characteristics of the athletes and explore potential relationships between variables. Data was collected through face-to-face sessions. Ethical approval was obtained from the International Islamic University Malaysia Research Ethics Committee (IIUM/504/14/11/2/IREC 2024-KAHS/DNS).

Participants

The participants of this study were selected among karate athletes who are training under the supervision of Majlis Sukan Pahang (MSP) and will be representing Pahang in the competitions. MSP is one of the government agencies that oversees the whole facilities and athletes that are chosen to train and compete for Pahang, including Karate. Athletes were included if they matched the criteria consisted of being around age 14-21 years old, registered under MSP, and competing at any level of the competition. Age 14-21 years old was considered to match the participation age for the Sukan Malaysia (SUKMA) competition. Athletes who were injured were excluded from this study.

The sample size of this study was determined using Krejcie and Morgan's sample size table (Krejcie, Morgan, 1970) as cited in Bukhari (2021). The population size of the karate were 35 and the expected sample size were 34 with 10% dropout rates were put into consideration. Purposive sampling was used in this study. The method was chosen to achieve the objective, which was very specific to only one group of the population. This purposive sampling technique followed strict inclusion and exclusion criteria for the sample. All coaches, parents, and guardians, including all other responsible parties such as the board of directors, were notified before the research was conducted.

Instruments

The assessment of this study was divided into three main parts, consisting of sociodemographic information, anthropometric profile, and sports nutrition knowledge. The sociodemographic questionnaire was adopted from the first part of the original NUKYA questionnaire by Espino et al., (2020) as cited in Nor Azizam et al., (2022). The question comprises age, gender, ethnicity, education level, entry to the state team, highest competition level, and

type of sport. Anthropometric profile of this study includes BMI, body fat percentage, and body somatotype. Meanwhile, knowledge of the athletes was assessed using the Nutritional Knowledge for Young Athletes (NUKYA) questionnaire.

Anthropometric Profile

i. BMI measurement

BMI measurement was calculated by taking the measurement of the athletes' height and body mass. Height was measured using the SECA 213 stadiometer; meanwhile, body mass was determined using the OMRON HBF-375 body fat analyzer (BIA). Both measurements were included in a formula, which is weight in kilograms (kg) divided by height in m^2 to obtain the BMI (Martinez-De-Quel et al., 2021; Villaseca-Vicuña et al., 2021).

ii. Body fat percentage

To calculate body fat percentage, body fat measurement was involved. Two sites' skinfolds were measured using a Harpenden skinfold calliper to calculate the body fat percentage, which includes triceps and subscapular. Both measurements were incorporated into the Slaughter et al., formula (Slaughter et al., 1988 as cited in Martinez-De-Quel et al., 2021):

$$BF\% = 1.33 \times \sum 2 - 0.013 \times \sum 2^2 - 2.5$$

where $\sum 2$ is the summation of two skinfolds (triceps and subscapular).

iii. Body somatotype

The somatotype or body type of an individual can be categorized into three main types, which are endomorph, mesomorph, and ectomorph. To determine the somatotype, 11 required measurements were taken as follows: body mass and body height, breadth of humerus and femur bioepicondylar, 4 skinfold measurements (triceps, subscapular, supraspinale, and calf), and 2 girths, which were arm (relaxed and extended) and calf. Skinfolds were taken at the right side of the participants using the Harpenden skinfold calliper. Meanwhile, circumference measurements were taken using Lufkin tape. Additionally, the diameter of the humerus and femur was determined using a vernier calliper. All the measurement procedures were referred to the International Society for the Advancement of Kinanthropometrics (ISAK).

Next, somatotype was determined using the Carter and Heath method (Carter, 2002) as summarized in the Table *1*. For endomorph, the height-corrected endomorph formula was used, where X is the summation of triceps, subscapular, and supraspinale skinfolds multiplied by 170.8 divided by height in cm(Carter, 2002). For mesomorph, Corrected Calf Girth (CCG) was used and subtracted from the skinfold in centimetres, and multiplied by pi from the calf girth. Lastly, for ectomorph decimal anthropometry depends on the Height-Weight Ratio (HWR) divided by the cube root of weight (stature/$mass^{1/3}$) (Carter, 2002).

Table 1 Decimal anthropometry formula

Somatotype	Formulae	Explanation
Mesomorph	0.858 x humerus breadth + 0.601 x femur breadth + 0.188 x corrected arm girth + 0.161 x corrected calf girth – height (0.131) + 4.5	Corrected Calf Girth (CCG) is obtained by subtracting the skinfold in centimetres and multiplying by pi from the calf girth. (Steward et al., 2002)
Endomorph	-0.7182 + 0.1451(X) – 0.00068(X^2) + 0.000014(X^3)	X – sum of triceps, subscapular, and supraspinale skinfold multiplied by (170.18/height in cm). This is a height-corrected endomorph.
Ectomorph	When HWR ≥ 40.75 Then; =0.732 HWR – 28.58 When 38.25<HWR<40.75 Then; 0.463 HWR – 17.63 When HWR ≤ 38.25 Then; =0.1	The formulae vary depending on the Height-Weight Ratio (HWR). HWR is height divided by the cube root of weight (stature/$mass^{1/3}$)

Next, the three-number somatotype rating was plotted on a 2-dimensional somatochart (*Figure 1*) by using X and Y coordinates using following formula(Carter, 2002; Slankamenac et al., 2021):

X = ectomorph – endomorph
Y = 2 X mesomorph – (endomorph + ectomorph)

Plotting the somatotype chart is important to determine the body somatotype of the athletes. It was used to identify the dominant body type of the athletes. Athletes can be classified into 13 categories. However, it can be further simplified into four larger groups:

1. Central: no component differs from another
2. Endomorph: dominant endomorphism, mesomorph, and ectomorph more than ½ unit lower
3. Mesomorph: dominant mesomorphism, endomorph, and ectomorph more than ½ unit lower
4. Ectomorph: dominant ectmorphism, endomorph, and mesomorph more than ½ unit lower (Carter, 2002)

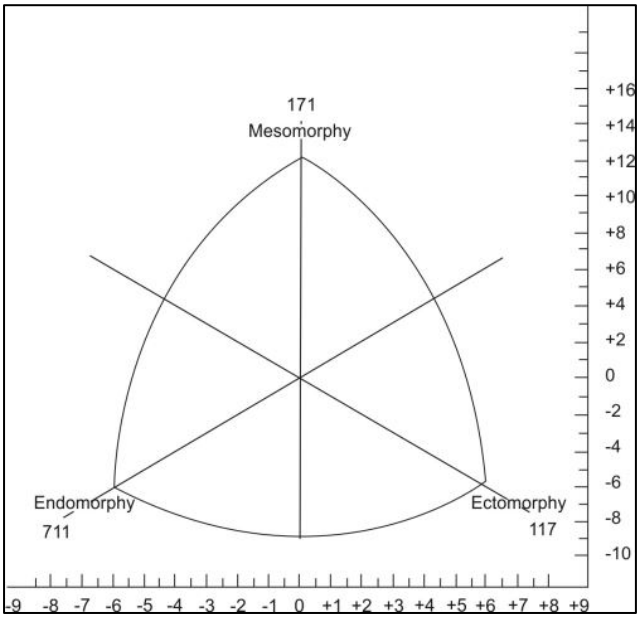


Figure 1 Example of Somatochart

Sports Nutrition Knowledge Questionnaire

The sports nutrition knowledge questionnaire was adopted from Azhar, (2023) to identify the level of sports nutrition knowledge of the athletes. The questionnaire includes 24 sets of questions with 57 items that cover 4 main sectors, which are macronutrients (27 items), micronutrients (19 items), hydration (8 items), and food periodicity (3 items). The maximum score was 57, and the minimum was 0. Participants who scored below 60% were classified as having inadequate sports nutrition knowledge, while those who scored 60% or higher were considered to have adequate knowledge (Nor Azizam et al., 2022).

Data Collection Procedure

Data collection was conducted in seven sessions. Athletes were grouped by gender, age, and competition level with assistance from the coaches. In each session, athletes alternated between completing the questionnaire and undergoing anthropometric measurements

Statistical analysis

Statistical analyses were conducted using IBM SPSS Statistics, version 26.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were applied to summarize sociodemographic data, with means and standard deviations reported for continuous variables, and frequencies and percentages for categorical variables. Associations between categorical variables were assessed using the Chi-square test, while Fisher's exact test was applied when expected cell counts were less than five. Spearman's rank-order correlation was used to examine relationships between continuous variables. All statistical tests were two-tailed, and a p-value of < 0.05 was considered statistically significant.

RESULTS

Table 2 presents the social demographic data of the 26 participants where the majority of them are males (84.6%), Malay (92.3%), with a mean age of 16.6 (±2.39) years old. Next, in terms of education level, most of the athletes were secondary school students (73.1%), and the remaining were in tertiary education.

Table 2 Sociodemographic information of the karate athletes (n=26)

Characteristics	n	%
Gender		
Male	22	84.6
Female	4	15.4
Race		
Malay	24	92.3
Indian	2	7.7
Education Level		
Secondary	19	73.1
Tertiary	7	26.9
Training level		
Preparation	10	38.5
Competition	16	61.5
Competition level		
State	3	11.5
National	8	30.8
International	15	57.7

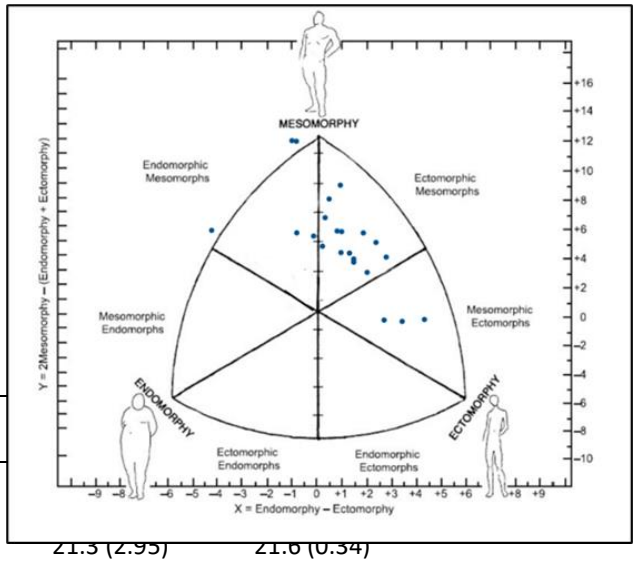
Ten athletes were in the preparation phase of the training (38.5%), while another 16 athletes (61.5%) were in the competition phase of the training. In addition, more than half of the athletes had competed at an international level as their highest competing level (57.7%). The final response rate was 72%, as several athletes were unable to participate due to scheduling constraints, competition commitments, and limited athlete availability during the data collection period.

Table 3 shows the anthropometric profile of the karate athletes according to their gender. For male athletes, the average of body mass index (BMI) was 21.3 (±2.95) kg/m^{2,} while the average value for body fat percentage was 12.0 (±4.86)%. Next, based on the calculation of body somatotype, male karate athletes tended to be predominantly mesomorphic (4.8 ±1.16), with moderate levels of endomorphy (1.6 ±0.91) and ectomorphy (2.7 ±1.22). The female karate athletes had a BMI of 21.6 (±0.34) kg/m^2 and a mean body fat percentage of 24.2 (±3.43) %. Somatotype analysis using the decimal anthropometry method showed mean scores of 4.1 ± 0.84 for endomorphy, 3.2 ± 1.31 for mesomorphy, and 3.0 ± 1.68 for ectomorphy, indicating a relatively balanced somatotype with a slight predominance of endomorphic characteristics.

Table 3 Anthropometric profile of the karate athletes (n=26)

Variable	**Mean (SD)**
Height (cm)	163.2 (6.38)
Weight (kg)	57.1 (9.28)
BMI (kg/m^2)	21.3 (2.71)
Body Fat Percentage	13.9 (6.4)
Skinfolds (mm)	
Triceps	7.2 (4.39)
Subscapular	7.9 (3.61)
Supraspinale	5.3 (3.35)
Calf	6.4 (7.20)
Breadths (cm)	
Hemerus	6.2 (0.39)
Femur	9.2 (0.50)
Girths (cm)	
Calf	34.5 (5.02)
Arm (Relaxed)	26.8 (3.15)
Arm (Tensed)	28.4 (2.85)
Somatotype	
Endomorphy	2.0 (1.26)
Mesomorphy	4.6 (1.31)
Ectomorphy	2.7 (1.27)
Somatochart coordinates	
X	0.7 (1.93)
Y	4.4 (3.90)

Next, Figure *2* illustrates the somatochart distribution of 22 male athletes. The majority (n = 14) were positioned in the ectomorphic-mesomorph region (upper right quadrant). Two athletes were located in the endomorphic-mesomorph region (upper left quadrant), while one athlete was placed on the left side of this region, and two just above it. Additionally, three athletes were distributed within the mesomorphic-ectomorph region. **Error! Reference source not found.** illustrates the somatochart distribution of four female athletes. Of these, one athlete was positioned in the endomorphic-ectomorph region (lower right quadrant), one within the mesomorphic-endomorph region (middle left quadrant), one within the ectomorphic-endomorph region, and another within the endomorphic-mesomorph region.



21.3 (2.95) 21.6 (0.34)
12.0 (4.86) 24.2 (3.43)

Figure 2 Male somatochart of karate athletes (n=22)

5.7 (2.80) 15.2 (2.14)
7.1 (3.02) 12.6 (3.23)

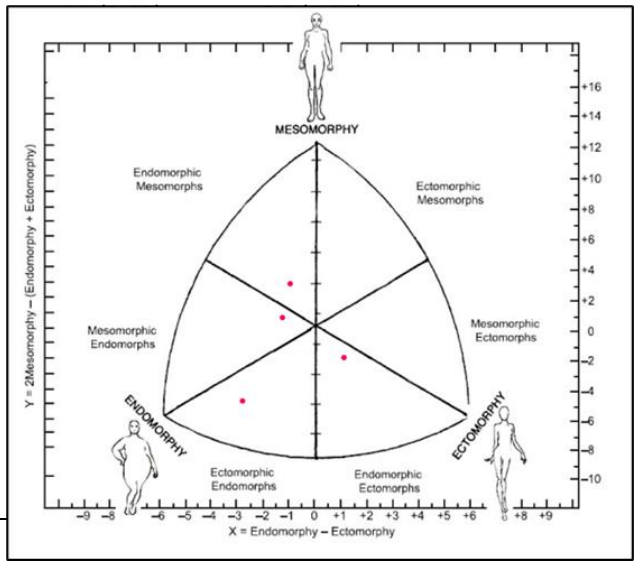


Figure 3 Female somatochart (n=4)

Table 4 presents the sports nutrition knowledge levels of the karate athletes. Overall, only 3 out of 26 athletes (11.5%) achieved a knowledge score above 60%, with a mean total score of 48.3 ± 10.01%. When stratified by gender, male athletes (n = 22) recorded a mean score of 46.5 ± 8.71%, while female athletes (n = 4) scored higher at 58.8 ± 11.59%. Knowledge was further assessed across seven subsections. The lowest mean score was observed in the fat knowledge category (43.2 ± 17.01%), whereas the highest was in food intake periodicity (65.4 ± 25.79%), where 18 athletes (69.2%) achieved scores above 60%.

Table 4 Sports Nutrition Knowledge (n=26)

Athletes	**Total score %**	**Adequate Knowledge Status (>60%)**	
	Mean (SD)	n	%
All (n=26)	48.3 (10.01)	3	11.5
Male (n=22)	46.5 (8.71)	1	4.5
Female (n=4)	58.8 (11.59)	2	50.0

The Fisher's exact test revealed no significant association between somatotype categories and sports nutrition knowledge adequacy among both male and female athletes (p = 1.000). Similarly, no significant correlation was observed between BMI and sports nutrition knowledge total score among male athletes (r = 0.002, p = 0.815) and between body fat percentage and sports nutrition knowledge total score among female athletes (r = –0.891, p = 0.200).

In contrast, a very strong negative correlation was found between BMI and sports nutrition knowledge total score among female athletes (r = –0.845, p = 0.0001) (Figure 4), indicating that athletes with higher BMI tended to have lower sports nutrition knowledge. Additionally, a moderate positive correlation was observed between body fat percentage and sports nutrition knowledge total score among male athletes (r = 0.427, p = 0.007) (Figure 5), suggesting that higher body fat percentage was associated with higher sports nutrition knowledge.

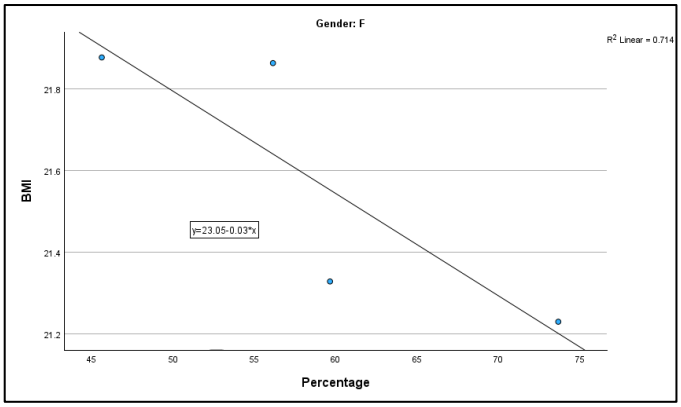


Figure 3 Association between Body Mass Index (BMI) and sports nutrition knowledge of female athletes.

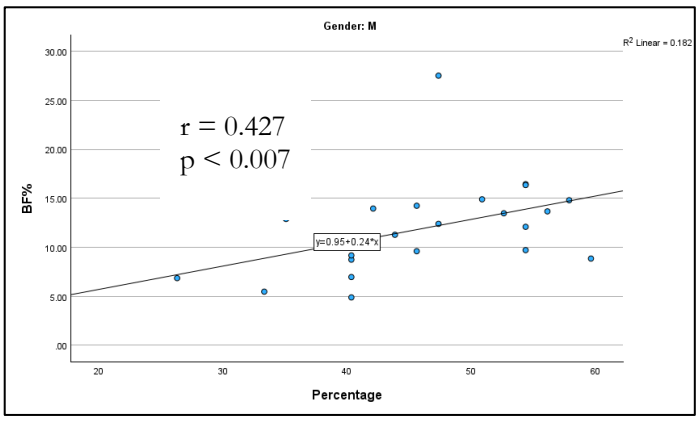


Figure 4 Association between body fat percentage and sports nutrition knowledge of male karate athletes.

DISCUSSION

This study identified the association between anthropometric profiles (body somatotype, body mass index [BMI], and body fat percentage) and sports nutrition knowledge (SNK) among karate athletes in Kuantan, Pahang. The findings demonstrated no significant association between body somatotype and SNK for both male and female athletes. With regard to BMI, no association was observed among male athletes, whereas a negative association was found among female athletes. For body fat percentage, a positive association with SNK was identified in male athletes, while no significant association was seen in female athletes. These results indicate that the relationship between anthropometry and SNK may differ by sex and by specific anthropometric indicators.

Anthropometric Profile

The anthropometric profile of the karate athletes in this study was evaluated through somatotype, BMI, and body fat percentage. Male athletes predominantly displayed an ectomorphic-mesomorphic somatotype, with only a few located in the endomorphic-mesomorphic region. This finding contrasts with a study by Slankamenac et al., (2021), which reported a predominant endomorphic-mesomorph somatotype among male karate athletes. Alkhayl et al. (2022) suggest that Asian populations may exhibit different upper-body strength and fat metabolism characteristics compared to Europeans, which may explain the contrast in the findings. Overall, the male karate athletes in this study exhibited a higher mesomorphic physique, indicating significant musculoskeletal robustness. This physique provides a muscle mass and strength advantage, beneficial for performance in combat sports (Tetiana et al., 2017).

Nevertheless, female athletes in this study exhibited greater variability, with four different somatotype classifications identified. Similar variations were reported by Bahamondes-Avil et al., (2023), who observed central, endomorphic-mesomorphic, and mesomorphic-endomorphic characteristics among their female karate athletes. Furthermore, Chaabene et al (2012) highlighted that elite-level female karatekas often exhibit mesomorphic-ectomorphic characteristics, emphasizing the importance of lower body fat and leaner profiles for optimal performance.

Regarding BMI, results revealed an average value of 21.3 (±2.95) kg/m² in male athletes and 21.6 (±0.34) kg/m² in female athletes. Compared to their non-athlete peers (Rezali et al., 2012), both groups demonstrated higher BMI, likely reflecting increased muscle mass. However, when comparing with other karate athletes, the male BMI values were lower than those reported by Martinez-De-Quel et al., (2021), which could be explained by differences in somatotype distribution. Female athletes, on the other hand, demonstrated BMI values consistent with previous studies, suggesting comparable trends across populations Martinez-De-Quel et al., 2021). While BMI is widely used, it does not differentiate between fat and lean mass, thus limiting its accuracy in reflecting body composition

In this study, male athletes exhibited a mean body fat percentage (BF%) of 12.0% (± 4.86). This is lower than the mean values previously reported for both Malaysian adolescent karateka (17.0% ± 12.8) (Aziz et al., 2025) and young male Spanish karatekas (21.3% ± 4.5) (Martinez-De-Quel et al., 2021). This comparative leanness accords with the predominantly ectomorphic–mesomorphic somatotype observed within the male sample of this study, which prioritizes musculoskeletal development over adiposity. In contrast, female athletes in the present study recorded a mean BF% of 24.2% (± 3.43). This value is substantially lower than the median BF% (~34.3%) observed in the Malaysian cohort reported by Aziz et al. (2025), and it is also lower than the mean BF% (27%) reported for young female Spanish karatekas (Martinez-De-Quel et al., 2021).

Taken together, the anthropometric findings suggest that both male and female karate athletes in this study generally demonstrated suitable body composition for competitive performance, though variations existed across somatotype, BMI, and body fat percentage. However, optimal anthropometric characteristics alone may not be sufficient to sustain performance without adequate sports nutrition knowledge, which plays a critical role in supporting training, recovery, and long-term health.

Sports Nutrition Knowledge

In this study, most karate athletes demonstrated inadequate levels of sports nutrition knowledge. This finding is consistent with reports among professional karate athletes in Ankara, Turkey, who also exhibited weak nutrition knowledge (mean score 34.18 ± 13.84) (Gobel, 2023). Such limited understanding is concerning, as established literature indicates that insufficient nutrition knowledge can negatively affect dietary quality, hinder recovery processes, and ultimately compromise overall athletic performance (Chaabène et al., 2012; Tabben et al., 2023). These results reinforce concerns that athletes may not be adequately equipped to translate nutritional requirements into practice, despite the high physical and metabolic demands of combat sports. Collectively, these findings highlight the urgent need for structured nutrition education and counselling programs tailored for karate athletes to enhance knowledge, improve dietary behaviors, and optimize performance outcomes.

Association of Anthropometric Profile and Sports Nutrition Knowledge of Karate Athletes.

The relationship between sports nutrition knowledge (SNK) and anthropometric indicators such as body fat percentage and somatotype is multifactorial and not always linear. In this study, a positive correlation between SNK and body fat percentage was observed among male athletes, indicating that greater nutrition knowledge did not necessarily

correspond to lower fat levels. This suggests a knowledge–practice gap, where athletes may possess adequate theoretical understanding but face challenges in consistently applying this knowledge to maintain optimal body composition.

The correlation between sports nutrition knowledge and body fat percentage is a complex subject that has been explored through various studies. Generally, higher sports nutrition knowledge correlates with better dietary habits and lower body fat percentage among athletes, although certain factors could potentially lead to higher body fat levels despite increased knowledge.

Research indicates that athletes with a strong understanding of sports nutrition often exhibit healthier dietary habits, positively influencing their body composition. For instance, Jagim et al., (2021) found that collegiate athletes with higher sports nutrition knowledge tended to have lower body fat percentages and fat mass, highlighting an inverse relationship between nutrition knowledge and body fat levels. This finding is echoed by Magee et al., (2023), who reported that athletes with greater nutrition knowledge were more likely to consume higher amounts of essential nutrients, such as carbohydrates and proteins, while minimizing unhealthy fats. This behavioral pattern supports maintaining a lower body fat percentage.

However, the relationship can be more intricate, especially when considering the possibility of higher body fat percentages despite strong nutrition knowledge. One contributing factor could be how athletes apply their nutritional knowledge in practice. Bakhtiar et al., (2021) noted that while older athletes often have more extensive nutrition training, their higher body mass index (BMI) is not directly correlated with their nutrition knowledge level, suggesting that age and other factors may influence body composition more than knowledge alone. Athletes may experience challenges in translating their knowledge into effective dietary practices, which can lead to continued accumulation of body fat.

Moreover, Staśkiewicz et al., (2023) emphasized that even with improved nutrition knowledge, complex determinants like lifestyle choices, metabolism variations, and psychological factors can hinder the achievement of optimal body fat levels. Even with an enhanced awareness of nutritional needs, athletes might still consume excess calories or fail to align their macronutrient distributions with performance demands, potentially resulting in higher fat mass.

Additionally, certain interventions aimed at improving nutrition knowledge do not always result in significant changes in dietary behaviors. For instance, Alvita Ghany et al., (2023) demonstrated that while nutrition education can enhance knowledge, it may not necessarily correlate with immediate behavioral changes regarding food choices. This disconnect can inadvertently lead to weight gain or the maintenance of higher body fat percentages among some athletes despite their understanding (Ghany et al., 2023). Such gaps highlight the complexity of the relationship between knowledge and practical application in nutritional habits.

Consequently, even as nutrition knowledge increases among athletes, it is essential to emphasize not only a theoretical understanding of dietary principles but also the comprehension of individual body needs, applying knowledge effectively, and integrating suitable lifestyle practices. This suggests that a high level of knowledge alone does not guarantee optimal body fat outcomes as knowledge does not consistently translate into healthier dietary behaviours due to various psychological, environmental, and lifestyle influences (Spronk et al., 2014; Contento, 2008; Heaney et al., 2011).

Consistent with this multifactorial relationship, the present study also found a significant negative correlation between sports nutrition knowledge and BMI among female athletes. This aligns with earlier findings showing that female athletes often demonstrate higher adherence to nutrition recommendations and place greater emphasis on weight management strategies compared to their male counterparts (Heaney et al., 2011; Hull et al., 2016). Research also suggests that female athletes tend to engage more actively in behaviour change when equipped with nutrition knowledge, resulting in more effective weight-control practices and healthier anthropometric profiles (Spronk et al., 2014). These gender-specific patterns indicate that the relationship between nutrition knowledge and anthropometric outcomes may differ by sex, underscoring the need for tailored nutrition education strategies

In contrast, no significant association was found between sports nutrition knowledge (SNK) and somatotype, suggesting that body type is influenced by factors beyond nutrition knowledge and less straightforward. Somatotype, which categorizes individuals based on body composition, often does not align as clearly with the level of nutritional understanding among athletes. One possible explanation for this disparity is that while nutrition knowledge can influence body composition, somatotype is also affected by genetic factors, training regimens, and overall body

structure that are independent of dietary practices (Penggalih et al., 2021). Research indicates that different sports typically attract individuals of varying somatotypes that fit the demands of the sport (e.g., endurance vs. strength sports) (González Macías & Flores, 2024). This specificity means that not all athletes with substantial nutritional knowledge will exhibit a somatotype corresponding to nutritional ideals recognized in sports.

Furthermore, while nutrition knowledge might positively influence dietary practices, several studies suggest that the actual impact on somatotype is influenced by broader factors, including lifestyle, genetics, and training methodologies, which may not be easily modified through education alone (Heaney et al., 2008; Janiczak et al., 2021). Indeed, barriers to optimal nutrition practices can exist despite elevated knowledge levels, as factors such as motivation, situational factors, and peer influence also significantly impact dietary choices among athletes (Ahmadi et al., 2022).

Strength and Limitation

This study is among the first to assess the association between anthropometric profiles and sports nutrition knowledge among Malaysian karate athletes, providing valuable baseline data for future research. The use of a validated and culturally adapted Nutrition Knowledge for Young Athletes questionnaire (Nor Azizam et al., 2022) and standardized ISAK anthropometric procedures strengthened the reliability and validity of the findings.

However, the small sample size particularly the limited number of female participants and the recruitment of athletes from only one region in Malaysia (Pahang) restricts the generalizability of the results. In addition, time constraints and the unavailability of advanced body composition tools may also have limited the depth of data collected. Despite these challenges, the study was conducted rigorously, and the findings remain meaningful within the context of Malaysian karate athletes.

CONCLUSION

This study revealed that while the karate athletes generally exhibit suitable physical characteristics for combat sports, their overall SNK levels are inadequate. Furthermore, a positive correlation between nutrition knowledge and body fat percentage was identified among male athletes only, whereas no significant association was observed between knowledge and somatotype in either gender. These findings suggest that greater nutrition knowledge does not necessarily translate into optimal body composition, highlighting a knowledge practice gap among athletes. Therefore, future interventions should focus not only on improving knowledge but also on promoting behavioral change and the practical application of nutrition principles. Expanding this research with larger and more diverse athlete samples may further clarify the influence of nutrition education on anthropometric and performance outcomes.

ACKNOWLEDGEMENT

The authors would like to express their sincere appreciation to all the karate athletes and coaches who participated and provided their time and cooperation throughout this study. Their support and commitment were invaluable in making this research possible.

REFERENCES

Abdul, S., Bukhari, R., & Ali, M. (2021). Sample Size Determination Using Krejcie and Morgan Table. https://doi.org/10.13140/RG.2.2.11445.19687

Ahmadi, F., Ebrahimi, M., & Kashani, V. (2022). Sports Nutritional Knowledge, Attitude, and Practice of Adolescent Athletes in Tehran, Iran. Asian Journal of Sports Medicine. https://doi.org/10.5812/asjsm-131584

Alkhayl, F. F. A., Ismail, A. D., Celis-Morales, C., Wilson, J., Radjenovic, A., Johnston, L., Welsh, P., Sattar, N., Gill, J. M. R., Preston, T., & Gray, S. R. (2022). Muscle protein synthesis and muscle/metabolic responses to resistance exercise training in South Asian and White European men. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-06446-7

Alvita Ghany, R. M., Hastuti, J., & Muslichah, R. (2023). Nutritional Status and Performance of Sepak Takraw Players Following Intervention by a Sports Dietitian. Journal of Health and Translational Medicine. https://doi.org/10.22452/jummec.vol26no2.18

Azhar, M. I. H. (2023). Sports nutrition knowledge and practice among adolescent athletes in Kuantan, Pahang (Unpublished master's thesis). International Islamic University Malaysia.

Aziz, M. A., Kunnabal, P., & Mohamed, N. A. (2025). Anthropometric and Physical Fitness Component Profiles in Malaysian Adolescence Recreational Karateka. Journal of Health Science and Medical

Research, 43(3), 1–10. https://doi.org/10.31584/jhsmr.20251142

Bakhtiar, M., Masud-ur-Rahman, Md., Kamruzzaman, Md., Sultana, N., & Rahman, S. S. (2021). Determinants of Nutrition Knowledge, Attitude and Practices of Adolescent Sports Trainee: A Cross-Sectional Study in Bangladesh. Heliyon. https://doi.org/10.1016/j.heliyon.2021.e06637

Bekele, B. A., Adane, A. K., & Kabiso, T. A. (2022). Anthropometric Characteristics Of Volleyball Players With Respect To Playing Positions: In Ethiopian Female Premier League. Central European Journal of Sport Sciences and Medicine, 39(3), 65–73. https://doi.org/10.18276/cej.2022.3-05

Carter, J. E. L. (2002). The Heath-Carter Anthropometric Somatotype-Instruction Manual-Somatotype Instruction Manual 2 Part 1: The Heath-Carter Anthropometric Somatotype-Instruction Manual. www.tep2000.com

Čaušević, D., Rani, B., Gasibat, Q., Čović, N., Alexe, C. I., Pavel, S. I., Burchel, L. O., & Alexe, D. I. (2023). Maturity-Related Variations in Morphology, Body Composition, and Somatotype Features among Young Male Football Players. Children, 10(4). https://doi.org/10.3390/children10040721

Contento I. R. (2008). Nutrition education: linking research, theory, and practice. Asia Pacific journal of clinical nutrition, 17 Suppl 1, 176–179.

Göbel, P. (2023). Nutrition knowledge levels and nutritional supplement beliefs of professional karate athletes. Journal of Health Sciences and Medicine, 6(1), 51-58. https://doi.org/10.32322/jhsm.1198884

González Macías, M. E., & Flores, J. (2024). Somatotype, Anthropometric Characteristics, Body Composition, and Global Flexibility Range in Artistic Gymnasts and Sport Hoop Athletes. Plos One. https://doi.org/10.1371/journal.pone.0312555

Heaney, S., O'Connor, H., Naughton, G., & Gifford, J. (2008). Towards an Understanding of the Barriers to Good Nutrition for Elite Athletes. International Journal of Sports Science & Coaching, 3(3), 391–401. https://doi.org/10.1260/174795408786238542

Heaney, S., O'Connor, H., Michael, S., Gifford, J., & Naughton, G. (2011). Nutrition knowledge in athletes: a systematic review. International journal of sport nutrition and exercise metabolism, 21(3), 248–261. https://doi.org/10.1123/ijsnem.21.3.248

Hull, M. V., Jagim, A. R., Oliver, J. M., Greenwood, M., Busteed, D. R., & Jones, M. T. (2016). Gender differences and access to a sports dietitian influence dietary habits of collegiate athletes. Journal of the International Society of Sports Nutrition, 13, 38. https://doi.org/10.1186/s12970-016-0149-4

Jagim, A. R., Fields, J. B., Magee, M. K., Kerksick, C. M., Luedke, J., Erickson, J. L., & Jones, M. T. (2021). The Influence of Sport Nutrition Knowledge on Body Composition and Perceptions of Dietary Requirements in Collegiate Athletes. Nutrients. https://doi.org/10.3390/nu13072239

Janiczak, A., Devlin, B. L., Forsyth, A., & Trakman, G. L. (2021). A Systematic Review Update of Athletes' Nutrition Knowledge and Association With Dietary Intake. British Journal of Nutrition. https://doi.org/10.1017/s0007114521004311

Kastrati, A., Gashi, N., Georgiev, G., Gontarev, S., & Kastrati, A. (2022). Somatotype Characteristics of Elite Young Athletes from the Republic of Kosovo. Sport Mont, 20(3), 47–52. https://doi.org/10.26773/smj.221008

Koropanovski, N., Berjan, B., Bozic, P. R., Pazin, N., Sanader, A., Jovanovic, S., & Jaric, S. (2011). Anthropometric and physical performance profiles of elite karate kumite and kata competitors. Journal of Human Kinetics, 30(1), 107–114. https://doi.org/10.2478/v10078-011-0078-x

Lijewski, M., Burdukiewicz, A., Stachoń, A., & Pietraszewska, J. (2021). Differences in anthropometric variables and muscle strength in relation to competitive level in male handball players. PLoS ONE, 16(12 December). https://doi.org/10.1371/journal.pone.0261141

Magee, M. K., Jones, M. T., Fields, J. B., Kresta, J., Khurelbaatar, C., Dodge, C., Merfeld, B., Ambrosius, A., Carpenter, M., & Jagim, A. R. (2023). Body Composition, Energy Availability, Risk of Eating Disorder, and Sport Nutrition Knowledge in Young Athletes. Nutrients. https://doi.org/10.3390/nu15061502

Martinez-De-Quel, O., Alegre, L. M., Castillo-García, A., & Ayán, C. (2021). Anthropometric and fitness normative values for young karatekas. Biology of Sport, 38(3),

351–357. https://doi.org/10.5114/BIOLSPORT.2021.99324

Nor Azizam, N. S., Yusof, S. N., Amon, J. J., Ahmad, A., Safii, N. S., & Jamil, N. A. (2022). Sports Nutrition and Food Knowledge among Malaysian University Athletes. Nutrients, 14(3). https://doi.org/10.3390/nu14030572

Penggalih, M. H. S. T., Solichah, K. M., Dewinta, M. C. N., Niamilah, I., Nadia, A., Pratiwi, D. E., … & Bactiar, N. (2021). Comparison of somatotype profiles and dietary intake of football athletes in different playing levels in indonesia. Jurnal Gizi Klinik Indonesia, 18(1), 18. https://doi.org/10.22146/ijcn.64653

Reid Reale. (2018). Acute Weight Management In Combat Sports: Pre Weigh-In Weight Loss, Post Weigh-In Recovery And Competition Nutrition Strategies. Gatorade Sports Science Institute , 31, 1–6.

Rezali, F. W., Chin, Y. S., & Yusof, B. N. M. (2012). Obesity-related behaviors of Malaysian adolescents: A sample from Kajang district of Selangor state. Nutrition Research and Practice, 6(5), 458–465. https://doi.org/10.4162/nrp.2012.6.5.458

Slankamenac, J., Bjelica, D., Jaksic, D., Trivic, T., Drapsin, M., Vujkov, S., Modric, T., Milosevic, Z., Drid, P., Campa, F., & Greco, G. (2021). Somatotype Profiles of Montenegrin Karatekas: An Observational Study. Public Health, 18, 12914. https://doi.org/10.3390/ijerph

Staśkiewicz, W., Grochowska-Niedworok, E., Zydek, G., Białek-Dratwa, A., Grajek, M., Jaruga-Sękowska, S., Kowalski, O., & Kardas, M. (2022). Changes in Body Composition During the Macrocycle of Professional Football Players in Relation to Sports Nutrition Knowledge. Frontiers in Nutrition. https://doi.org/10.3389/fnut.2022.981894

Spronk, I., Kullen, C., Burdon, C., & O'Connor, H. (2014). Relationship between nutrition knowledge and dietary intake. The British journal of nutrition, 111(10), 1713–1726. https://doi.org/10.1017/S0007114514000087

Staśkiewicz, W., Grochowska-Niedworok, E., Zydek, G., Grajek, M., Krupa-Kotara, K., Białek-Dratwa, A., Jaruga-Sękowska, S., Kowalski, O., & Kardas, M. (2023). The Assessment of Body Composition and Nutritional Awareness of Football Players According to Age. Nutrients. https://doi.org/10.3390/nu15030705

Stewart, A. D., Stewart, A., & Reid, D. M. (2002). Correcting calf girth discriminates the incidence of falling but not bone mass by broadband ultrasound attenuation in elderly female subjects. Bone, 31(1), 195–198. https://doi.org/10.1016/s8756-3282(02)00811-6.

Sunuwar, D. R., Singh, D. R., Bohara, M. P., Shrestha, V., Karki, K. B., & Singh Pradhan, P. M. (2022). Association of Nutrition Knowledge, Practice, Supplement Use, and Nutrient Intake With Strength Performance Among Taekwondo Players in Nepal. Frontiers in Nutrition. https://doi.org/10.3389/fnut.2022.1004288

Thapa, M., Neupane, A., Shrestha, S. K. D., Nepal, P., Upadhyaya, A., Niraula, P., Shrestha, R. K., & Sunuwar, D. R. (2023). Factors affecting nutritional knowledge, attitude, practices and dietary intake among national players in Kathmandu, Nepal: a cross-sectional study. BMC Sports Science, Medicine and Rehabilitation, 15(1). https://doi.org/10.1186/s13102-023-00691-7

Villaseca-Vicuña, R., Molina-Sotomayor, E., Zabaloy, S., & Gonzalez-Jurado, J. A. (2021). Anthropometric profile and physical fitness performance comparison by game position in the Chile women's senior national football team. Applied Sciences (Switzerland), 11(5), 1–16. https://doi.org/10.3390/app11052004Arksey,

ORIGINAL ARTICLE

A Comparison of Knowledge, Attitude and Practices Related to Milk and Dairy Products among Male and Female Orphanage Adolescents in Kuantan, Pahang

Siti Nur Amirah Sheikh Hishamuddin[1], Nur Syazwani Azmi[2], Aliza Haslinda Hamirudin[2,3*]

[1]Faculty of Health, University of Cyberjaya, Persiaran Bestari, Cyber 11, 63000, Cyberjaya, Selangor, Malaysia
[2]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[3]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia

Keywords:
milk and dairy products; adolescents; knowledge; attitude; practice

ABSTRACT

Background: Adequate milk and dairy product intake throughout life is essential, particularly among adolescents, for their growth and bone health due to dietary calcium content. Nevertheless, research regarding knowledge, attitude, and practices related to milk and dairy products among Malaysian orphanage adolescents according to gender is scarce. Hence, this study aimed to determine the knowledge, attitude, and practices towards milk and dairy products among male and female residents of orphanage institutions. **Methods:** In this cross-sectional study, seventy (n=70) secondary school-age adolescents (13-17 years old) were recruited through convenience sampling from five (n=5) orphanage institutions in Kuantan, Pahang. Data were collected using a self-administered questionnaire that consisted of sociodemographic information, knowledge, attitude, and practices of milk and dairy products. The data were analysed with SPSS software using chi-square and independent t-tests, at the statistically significant level of $p<0.05$. **Results:** The mean (± SD) knowledge score was 7.87 ± 1.23. Most respondents were categorized in the good knowledge category with the percentage of 70% (n=49), 25.7% (n=18) in the moderate category, and only 4.3% (n=3) of them scored in the poor category. Positive attitudes were demonstrated in more than 50% respondents, although a misconception that milk contributes to weight gain was identified. About 48.5% (n=34) of participants drink milk every day, but only 21.4 % of subjects (n=15), or approximately only one in five of the respondents, consumed two glasses of milk per day. No significant gender differences were identified for knowledge, attitude, and practices. **Conclusion: The** Majority of adolescents in orphanage institutions showed good knowledge and positive attitudes towards milk and dairy products, but lower intake than recommended, with no differences by gender. Nutrition education and institutional supports are warranted to enhance dairy intake for optimal growth in this population.

INTRODUCTION

Milk and dairy products are widely recognized as essential components of a balanced diet due to their high nutritional value and contribution to human growth and health. The consumption of milk has increased over the past few decades as studies show that milk is rich in nutrients (Kourkouta Lambrini et al., 2020; Givens., 2020). Besides, they have high-quality protein, calcium, phosphorus, vitamin D, and vitamins B that are important for bone development, muscle function, and metabolic process (Geller et al., 2022). The Malaysian Dietary Guidelines (MDG) and Recommended Nutrient Intake (RNI) for Malaysia highlighted that the inclusion of milk and dairy products is needed as part of daily intake (RNI, 2017; MDG, 2020). Based on the requirements, adolescents need to consume 1300 mg of calcium daily. Despite these recommendations, the consumption of milk and dairy products is still inadequate among adolescents in Malaysia (Sallehuddin et al., 2021; Shlisky et al., 2022).

Furthermore, children and adolescents who are living in orphanage institutions are more vulnerable than the population in terms of nutritional health (DeLacey et al., 2020; Shaziman et al., 2017). They are at high risk of malnutrition, impaired intellectual growth in the early years, and loss of family attachment. A study in Lahore, Pakistan, highlighted that the orphaned adolescents had subclinical deficiencies due to improper dietary habits (Khan et al., 2024). Besides, orphaned children are highly susceptible to inadequate nutrition due to factors like insufficient care, inadequate hygiene, and limited awareness (Safa et al., 2024).

* *Corresponding author.*
E-mail address: aliza@iium.edu.my

Nutrition is one of the important aspects in the development of adolescents. Therefore, assessing knowledge, attitudes, and practices (KAP) related to food and nutrition provides valuable insights into how individuals understand and apply nutritional information in their daily lives (Marías & Glasauer, 2014). From the previous study, one of the factors that is related to milk and dairy consumption is insufficient nutritional knowledge (Hamulka et al., 2024). Hence, another study revealed that the consumption of dairy products was higher with good nutritional knowledge (Mashallah Mohamed Abdul-Aziz et al., 2021). In Malaysia, a cross-sectional survey among secondary school students showed that only 23.2% of adolescents reported consuming milk or milk products at least twice daily over the past 30 days (Salehuddin et al., 2022). Furthermore, analysis of the South East Asian Nutrition Surveys (SEANUTS) data showed that the prevalence of vitamin D insufficiency was lower among dairy consumers compared to non-dairy consumers, 39.4% and 53.8%, respectively (Bao et al., 2018).

While previous studies have explored KAP related to milk and dairy consumption among the general adolescent population, this study aimed to focus on adolescents living in orphanage institutions. Despite the various causes of insufficient intake, health behaviours may be influenced by various factors, including knowledge and attitude (Alves, 2024). To date, no study has been conducted in Malaysia to compare KAP related to milk and dairy products among male and female residents of orphanage institutions. Therefore, this study aimed to compare KAP related to milk and dairy products among residents of orphanage institutions in Kuantan, Pahang, according to gender.

MATERIALS AND METHODS

Study Design and Population

A cross-sectional study was conducted among orphanage residents in five selected orphanage institutions in Kuantan, Pahang. The participants were selected based on inclusion criteria, which were male and female adolescents aged 13 to 18 years old from the institutions and were free from physical and mental disabilities. Meanwhile, participants who had milk and dairy product intolerance or allergy, or who had any medical illness were excluded from this study.

Sampling Method

The estimation of sample size was based on an expected prevalence of inadequate dietary calcium intake among adolescents (Chee et al., 2008), with 80% power and a 7% margin of error. After accounting for a 10% non-response rate, the final minimum required sample size was 35 adolescents per group. Hence, a total of seventy (n=70) respondents were required for comparison between gender (n=35 male, n=35 female).

A convenient sampling method was conducted for the recruitment of the participants. Thus, participants who fulfilled the inclusion criteria were provided with an informed consent form to fill in for their voluntary participation. Next, a self-administered questionnaire was distributed. The questionnaire consisted of two main parts: (1) demographic information that included gender, year of study in secondary school, and race, health status, and (2) questions assessing knowledge, attitudes, and practices related to milk and dairy products among residents of orphanage institutions in Kuantan, Pahang. The questionnaire was adopted from Millen et al. (2005), and a pilot test was performed before the study's conduct. For knowledge, a total of nine questions were asked regarding milk and dairy products. The respondents received one mark for each correct answer and 0 marks for a wrong answer. The knowledge score was classified into three good categories (≥8 or above), moderate (6 -7), and poor (≤5).

Data Analysis

All data were analyzed using SPSS version 12.0. Descriptive statistics were used to analyze demographic information. Meanwhile, knowledge, attitudes, and practices related to milk and dairy products among female and male residents of orphanage institutions in Kuantan, Pahang were examined using independent t-test and a chi-square test. P value <0.05 was set as statistically significant.

Ethics Application

Ethics approval was obtained from the International Islamic University Malaysia (IIUM) Research Ethics Committee (IREC), IREC NO: IREC 2018-185.

RESULTS

Sociodemographic Characteristics

A total of seventy (n=70) male and female residents of orphanage institutions from Kuantan, Pahang, participated in this study. The respondents' age ranged between 13 to 17 years old, with the mean (±SD) of 14.29 ± 2.04 years old. Table 1 shows the sociodemographic characteristics of the study participants.

Table 1: Sociodemographic characteristics of participants (n=70)

Variables	n	%
Form (in Secondary School)		
Form 1	23	32.9
Form 2	19	27.1
Form 3	10	14.3
Form 4	9	12.9
Form 5	9	12.9
Gender		
Male	35	50
Female	35	50
Race		
Malay	69	98.6
Indian	1	1.4

The majority of participants were Form 1 (32.9%), followed by Form 2 (27.1%), Form 3 (14.3%), Form 4 (12.9%) and Form 5 (12.9%) respectively. 98.6% (n=69) respondents were Malay, while only 1.4% (n=1) was Indian participant.

Knowledge, Attitude and Practice regarding Milk and Dairy Products

Knowledge

The mean (± SD) knowledge score regarding milk and dairy products of all participants was 7.87 ± 1.23. Most respondents were categorized in the good knowledge category, with the percentage of 70% (n=49), 25.7% (n=18) of them in the moderate category, and only 4.3% (n=3) in the poor category. Table 2 presents the knowledge score towards milk and dairy products according to gender.

Table 2: Knowledge score towards milk and dairy products among male and female participants

Variable	Gender	
	Male (n=35) n (%)	**Female (n=35) n (%)**
Good for health		
Yes	32 (91.4)	35 (100)
No	3 (8.6)	0 (0)
High in calcium		
Yes	34 (97.1)	33 (94.2)
No	1 (2.8)	2 (5.7)
Vital for bone, teeth, brain		
Yes	33 (94.2)	34 (97.1)
No	2 (5.7)	1 (2.8)
Recommendation to drink 2 glasses/day		
Yes	26 (74.2)	27 (77.1)
No	9 (25.8)	8 (22.9)
Vital for children growth		
Yes	33 (94.2)	35 (100)
No	2 (5.7)	0 (0)
Contribute to being fat		
Yes	4 (11.4)	2 (5.7)
No	31 (88.6)	33 (94.2)
Only drink when sick		
Yes	5 (14.2)	1 (2.8)
No	30 (85.7)	34 (97.1)
In the food pyramid, milk and dairy products should be consumed moderately		
Yes	25 (71.4)	25 (71.4)
No	10 (28.6)	10 (28.6)
Milk should be consumed throughout life, not only during childhood		
Yes	23 (65.7)	29 (82.8)
No	12 (34.2)	6 (17.1)

An independent t-test was conducted to compare knowledge of milk and dairy products between males and females. There was no significant difference (p=0.064) for knowledge score on milk and dairy products between male and female adolescents (Table 3). The mean score for male is 7.60± 1.42 while the mean score for females is 8.14± 0.94.

Table 3: Total knowledge score of milk and dairy products among male and female participants

Gender	n	Mean	SD	Mean difference (96% CI)	t statistics	df	p-value[a]
Male	35	7.60	1.42	-0.543	-1.885	68	0.064
Female	35	8.14	0.94				

Attitude

Table 4 shows the respondents' attitude regarding milk and dairy products. Majority of the respondents perceived milk as tasty (90%, n=63). For this question, 85.7% (n=30) male answered 'yes' and 94.2% (n=33) female answered 'yes'. The results demonstrated that 58.5% (n=41) would read food labels when buying milk and dairy products. 64.2% (n=45) would drink milk without instructions from

fellows. Most male (74.2%, n=26) and female (68.5%, n=24) respondents responded 'yes' to encouragement from social media. Nevertheless, no statistically significant findings were identified for all attitude towards milk and dairy products between male and female respondents.

Table 4: Attitude towards milk and dairy products among male and female participants

Variable	Gender		Chi-square X^2	p-value[a]
	Male (n=35) n (%)	Female (n=35) n (%)		
Tasty and delicious				
Yes	30 (85.7)	33 (94.2)	1.429	0.232
No	5 (14.2)	2 (5.7)		
Price is affordable				
Yes	23 (65.7)	30 (85.7)	3.807	0.051
No	12 (34.2)	5 (14.2)		
Suggestion from fellows to drink milk				
Yes	21 (60)	25 (71.4)	1.014	0.314
No	14 (40)	10 (28.6)		
Read the label when buying milk and dairy products				
Yes	17 (48.5)	24 (68.5)	2.885	0.089
No	18 (51.5)	11 (31.5)		
Drink without any instruction from fellows				
Yes	25 (71.4)	20 (57.1)	1.556	0.212
No	10 (28.6)	15 (42.9)		
Encouragement from social media				
Yes	26 (74.2)	24 (68.5)	0.280	0.597
No	9 (25.8)	11 (31.5)		

[a]Chi-square test

Practice

Table 5 shows the practice towards milk, yogurt, and cheese consumption among orphanage residents. Results showed that the majority of respondents, 84.3% (n=59) like to drink milk, of whom 44.3% (n=31) were male and 40% (n=28) were female, respectively.

The majority of the respondents like to eat yogurt (82.8%, n=58) which represents 77.1% male and 88.5% female, respectively. For the type of yogurt, the distribution of preferences of full cream yogurt and low-fat yogurt from all respondents was similar which is 32.8% (n=23) respectively, followed by free fat yogurt 14.2% (n=10) and 17.1% (n=12) not eating yogurt.

For cheese preferences, 71.4% of respondents like to eat cheese, 41.4% (n=29) were male, and 30% (n=21) were female. Regarding the type of cheese, low-fat cheese is the most preferred by female participants, whilst most males prefer cheddar cheese.

Table 5: Milk, yogurt, and cheese consumption among male and female participants (n=70)

Variables	Gender	
	Male (n=35) n (%)	Female (n=35) n (%)
MILK		
Like to drink milk		
Yes	31 (88.5)	28 (80)
No	4 (11.5)	7 (20)
Flavour of milk		
Full cream milk	4 (11.4)	6 (17.1)
Low-fat milk	3 (8.5)	1 (2.8)
Strawberry	7 (20)	6 (17.1)
Chocolate	14 (40)	18 (51.4)
Coffee	3 (8.5)	0 (0)
Green tea	1 (2.8)	1 (2.8)
Honey	2 (5.7)	1 (2.8)
Corn	0 (0)	0 (0)
Others	0 (0)	0 (0)
Not drinking milk	1 (2.8)	2 (5.7)
YOGURT		
Like to eat yogurt		
Yes	27 (77.1)	31 (88.5)
No	8 (22.8)	4 (11.5)
Type of yogurt		
Full cream	12 (34.2)	11 (31.4)
Low fat	6 (17.1)	17 (48.5)
Free fat	7 (20)	3 (8.5)
Others	2 (5.71)	0 (0)
Not eating yogurt	8 (22.8)	4 (11.4)
Flavour of yogurt		
Original	3 (8.5)	3 (8.5)
Strawberry	9 (25.7)	9 (25.7)
Blackcurrant	2 (5.7)	6 (17.1)
Mixed fruits	7 (20)	11 (31.4)
Kiwi	1 (2.8)	0 (0)
Mango	5 (14.2)	2 (5.7)
Not eating yogurt	8 (22.8)	4 (11.4)

CHEESE		
Like to eat cheese		
Yes	29 (82.8)	21 (60)
No	6 (17.2)	14 (40)
Type of cheese		
Cheddar	14 (40)	10 (28.5)
Low fat	13 (37.1)	13 (37.1)
Less salt	5 (14.2)	6 (17.1)

Table 6 demonstrates intake of milk according to gender, comprised of 80% (n=28) males and 62.8% (n=22) females, respectively. Moreover, 48.5% consumed at least 1 serving of milk per day, and only 21.4% (n=15) or one out of five respondents achieved the recommendation of taking two glasses of milk daily. Overall, 28.5% (n=20) of the adolescents reported not drinking milk at all. Non-consumption was higher among females (18.5%), and more than adequate consumption (>2 glasses) was higher among males (7.14%). Nevertheless, no statistically significant difference in practices was found between males and females.

Table 6: Frequencies of milk, yogurt, and cheese intake among male and female participants

Variables	**Gender**			
	Male n (%)	**Female n (%)**	**Chi square X^2**	**p-value[a]**
MILK				
Daily milk consumption (1 glass= 250 ml)				
>6 times	0 (0)	0 (0)	7.901	0.095
≤ 6 times	28 (80)	22 (62.8)		
YOGURT				
Daily yogurt consumption (1 small container = 125 g)				
>6 times	0 (0)	0 (0)	1.619	0.445
≤ 6 times	17 (48.5)	17 (48.5)		
CHEESE				
Daily cheese consumption				
>6 slices	0 (0)	0 (0)	4.990	0.172
≤ 6 slices	14 (40)	14 (40)		

[a]Chi-square test

DISCUSSION

This study aimed to compare knowledge, attitude, and practice related to milk and dairy products among male and female orphanage adolescents in Kuantan, Pahang.

Knowledge

Our study demonstrated that most adolescents from orphanage institutions were categorized in the good knowledge category, whilst the mean knowledge is in the moderate category. In contrast, Shaziman et al. (2017) reported poor nutritional knowledge among the majority of orphanage residents in Malacca and Selangor. The nutritional education provided in the orphanage settings could help improve their knowledge, which could subsequently improve their practice on milk and dairy products (Pysz et al. 2015). Milosavljević et al (2015) highlighted that knowledge regarding food or beverages is particularly important to health as it is associated with eating behaviour, including dietary intake, frequency of intake, and preference. However, understanding of this topic is still scarce, especially considering the adolescents' living environments in orphanage institutions and their influence on their intake.

In addition, the present study found no significant mean difference in knowledge of milk and dairy products between male and female respondents. Other studies among adolescents identified contrasting findings in which higher nutritional knowledge was identified among females than males (Naeeni et al. 2014, Tallarini et al. 2014). For the question 'milk contributes to being fat', the result identified that more than half respondents disagreed. In contrast to a survey by Stice et al. (2011), they demonstrated that children's concerns about a healthy diet were related to body image rather than diet-related chronic diseases. Despite this, other studies reported that there is a misconception that milk is fattening, which may lead children, especially females, to limit their milk intake to lose or maintain weight (Dror 2014, Censi et al. 2020).

Besides, participants in the present study responded that milk is good for humans. This result showed that most of the adolescents were aware of the benefits of milk for them. This is also aligned with Bus & Worsley (2003) findings that the participants thought that drinking milk makes them healthy, and taste was an important characteristic for their milk choice.

Attitude

Our findings demonstrated that the majority of respondents had positive attitudes towards milk and dairy products, although a misconception that milk contributes to weight gain was identified. This is supported by Louie et al. (2011) demonstrated that weight-related outcomes can influence attitudes towards milk among young individuals. Besides, Shaziman et al. (2017) demonstrated that 67.1% of adolescents living in orphanage institutions had a good nutritional attitude. Larson et al. (2006) identified that males had a higher taste preference for milk than females; whilst this preference in both genders showed a significant positive correlation with greater calcium intake. Nevertheless, our study found no significant differences for all attitude components between male and female respondents; related to reading the food labels when buying milk and dairy products, suggestions and instructions from caretakers to drink milk, and encouragement from social media.

Taste appears to be an important factor influencing adolescents' food choices (Giskes et al., 2005), including their decision to drink milk (Racey et al., 2017). From this, it can be concluded that milk flavour will influence adolescents' milk drinking behaviour. This is in line with a study conducted by Bus & Worsley (2003) in which dairy milk is considered to have good sensory properties by the majority of respondents.

Parental influence is an important factor in shaping adolescents' attitudes toward milk and dairy products (Schiano et al., 2022). In this case, caretakers in the orphanage institutions could play a significant role in enhancing adolescents' attitudes, which can subsequently contribute to adequate intake. Moreover, school-based intervention can positively influence adolescents' choice of healthier dairy product options (Tavares Filho et al., 2025). Another factor that influences milk intake is affordability. Although our present study did not find differences between genders, most respondents agreed that the price of milk and dairy products is affordable. Taheri et al. (2017) demonstrated that the choice of dairy products among Malaysians is affected by price, along with the quality and brand of the products.

Practice

Our findings indicated that most respondents like to drink milk, whilst 48.5% consumed at least 1 serving of milk per day. Although it is likely that other type of dairy products was consumed from other sources, it is evident that consumption of milk as a beverage was suboptimal. This corresponds to attitudes and beliefs about a particular food predict individuals' consumption (Wham & Worsley, 2003, Shaziman et al., 2017). There may be knowledge and attitude factors underlying inadequate milk consumption that require attention, in which nutrition education programs are beneficial to promote more healthful behaviours (Wang et al., 2015).

Malaysia Dietary Guideline for Children and Adolescents (2013) indicated that adolescents should consume three servings of milk and dairy products per day, whilst the updated Malaysia Dietary Guideline for Children and Adolescents (2023) recommended 2 servings per day for female and 3 servings per day for male adolescents respectively. However, only one in five (21.4%) of our respondents achieved the intake of two glasses of milk daily. Non-milk drinkers were higher among females, whilst more male respondents consumed at least two glasses of milk daily. Noteworthy, our finding is parallel with the National Health Morbidity Survey in 2022 among Malaysian adolescents that reported 23.2% milk or milk products consumption of two times per day within the previous 30 days, with no significant intake difference between gender (Institute for Public Health, 2022).

Our study also identified that there was no significant difference in practice towards milk and dairy products between males and females. Contrary finding was reported by Larson et al. (2006) as male adolescents have higher consumption than females. An Iranian study demonstrated that being male is a significant contributor to low milk and dairy product intake (Shokrvash et al. 2015). The authors also highlighted that female adolescents have significantly reached the recommended three servings of recommended milk and dairy products intake compared to males.

Availability of milk at home (Arcan et al., 2007) or schools (Lee et al., 2019) was consistently associated with higher milk intake. According to Hanson et al (2005), increased availability of milk and dairy products in the household can be a contributor to promoting the consumption, which will replace or eliminate the consumption of carbonated drinks among adolescents. In this present study, the respondents stay at orphanage institutions. Thus, most of their daily dietary intake is based on the prepared meals in the institutions, which play a vital role in meal provision.

Limitations and Strengths

Limitations of this study include that the study was conducted only in five selected orphanage institutions in Kuantan, Pahang. This would limit the generalizability of the results of this study. Other than that, the data were self-reported. Potential errors from memory and

estimation of portion size of milk, yogurt and cheese intake may occur due to children's cognitive ability to record or remember their diets. Despite of that, the strength of this study is being at the forefront in comparing knowledge, attitude and practices of milk and dairy products consumption among orphanage adolescents by gender.

CONCLUSION

In conclusion, the present study provides insights into the knowledge, attitude, and practice of milk and dairy products consumption among the residents of orphanage institutions in Kuantan, Pahang, with no significant differences identified between male and female adolescents. Milk intake among most respondents was lower than the recommendation for Malaysian adolescents, although the majority of them have good knowledge and attitude.

For future research, a larger sample size that incorporates respondents from orphanage institutions across various states, rural and urban regions is recommended. Besides, dietary aid tools can be used to facilitate participants in answering the questionnaire, such as using household measurements and pictorial design to estimate their intake. Nutrition intervention in this population and institutional support is warranted to improve their knowledge, attitudes, and practices of milk and dairy products.

ACKNOWLEDGEMENT

Our sincere gratitude to the orphanage institutions that provided approval for the conduct of the study and to the respondents for their participation.

REFERENCES

Alves, R. (2024). The relationship between health-related knowledge and attitudes and health risk behaviours among Portuguese university students. Global Health Promotion, 31(1), 36–44. https://doi.org/10.1177/17579759231195561

Arcan, C., Neumark-Sztainer, D., Hannan, P., Van Den Berg, P., Story, M., & Larson, N. (2007). Parental eating behaviours, home food environment and adolescent intakes of fruits, vegetables and dairy foods: longitudinal findings from Project EAT. Public health nutrition, https://doi.org/10.1017/s1368980007687151

Bao, K. L. N., Sandjaja, S., Poh, B. K., Rojroongwasinkul, N., Huu, C. N., Sumedi, E., Aini, J. N., Senaprom, S., Deurenberg, P., Bragt, M., & Khouw, I. (2018). The consumption of dairy and its association with nutritional status in the south east Asian nutrition surveys (SEANUTS). Nutrients, 10(6). https://doi.org/10.3390/nu10060759

Bus, A. M. and Worsley, A. (2003). Consumers' sensory and nutritional perceptions of three types of milk. Public Health Nutrition, 6(2), 201-208. https://doi.org/10.1079/phn2002417

Censi, S., Manso, J., Barollo, S., Mondin, A., Bertazza, L., Marchi, M. D., … & Services, o. b. o. t. F. a. N. H. (2020). Changing dietary habits in veneto region over two decades: still a long road to go to reach an iodine-sufficient status. Nutrients, 12(8), 2399. https://doi.org/10.3390/nu12082399

Chee, Y. F., Roseline Yap, W. K., & Siti Sabariah, B. (2008). Weight Status and Dietary Intake among Female Children and Adolescents Aged 6-17 Years in a Welfare Home, Kuala Lumpur. Malaysian Journal of Nutrition, 14(1), 79–89.

DeLacey, E., Tann, C., Groce, N., Kett, M., Quiring, M., Bergman, E., Garcia, C., & Kerac, M. (2020). The nutritional status of children living within institutionalized care: A systematic review. In PeerJ (Vol. 8). PeerJ Inc. https://doi.org/10.7717/peerj.8484

Dror, D. K. (2014). Dairy consumption and pre-school, school-age and adolescent obesity in developed countries: a systematic review and meta-analysis. Obesity Reviews, 15(6), 516-527. https://doi.org/10.1111/obr.12158

Geller, S. G., Clark, B. E., Pope, L., Niles, M. T., & Belarmino, E. H. (2022). Investigating Knowledge on Calcium and Preferences for Dairy vs. Plant-Based Alternatives. In Journal of Healthy Eating and Active Living (Vol. 2022, Issue 2).

Giskes, K., Patterson, C., Turrell, G., & Newman, B. (2005). Health and nutrition beliefs and perceptions of brisbane adolescents. Nutrition & Dietetics, 62(2-3), 69-75. https://doi.org/10.1111/j.1747-0080.2005.00002.x

Hamulka, J., Czarniecka-Skubina, E., Górnicka, M., Gębski, J., Leszczyńska, T., & Gutkowska, K. (2024). What Determinants Are Related to Milk and Dairy Product Consumption Frequency among Children Aged 10–12 Years in Poland? Nationwide Cross-Sectional Study.

Nutrients , 16(16). https://doi.org/10.3390/nu16162654

Hanson, N. I., Neumark-Sztainer, D., Eisenberg, M. E., Story, M., & Wall, M. (2005). Associations between parental report of the home food environment and adolescent intakes of fruits, vegetables and dairy foods. Public health nutrition, 8(1), 77-85.

Institute for Public Health (IPH) 2022. Technical Report National Health and Morbidity Survey (NHMS) 2022: Adolescent Health Survey, Malaysia

Khan, H., Sohail, Z., Khan, N., Zahra, S., Niazii, R. K., Anwaar, U., Zaheer, A., Moazzam, M., & Arshad, D. (2024). Assessment of the Nutritional Status of Adolescents in Orphanage from Lahore, Pakistan. International Health Review, 3(2), 43–59. https://doi.org/10.32350/ihr.32.04

Kourkouta Lambrini, Frantzana Aikaterini, Koukourikos Konstantinos, Iliadis Christos, Papathanasiou V. Ioanna, & Tsaloglidou Areti. (2020). Milk Nutritional Composition and Its Role in Human Health. Journal of Pharmacy and Pharmacology, 9(1). https://doi.org/10.17265/2328-2150/2021.01.002

Larson, N. I., Story, M., Wall, M., & Neumark-Sztainer, D. (2006). Calcium and dairy intakes of adolescents are associated with their home environment, taste preferences, personal health beliefs, and meal patterns. Journal of the American dietetic association, 106(11), 1816-1824.

Louie, J. C. Y., Flood, V., Hector, D., Rangan, A., & Gill, T. P. (2011). Dairy consumption and overweight and obesity: a systematic review of prospective cohort studies. Obesity Reviews, 12(7). https://doi.org/10.1111/j.1467-789x.2011.00881.x

Macías, Y.F., & Glasauer, P. (2014). Guidelines for assessing nutrition-related knowledge, attitudes and practices : KAP manual. Food and Agriculture Organization of the United Nations.

Mashallah Mohamed Abdul-Aziz, Aya Abdul-Salam Algomati, Timh Salem Alhasi, Mailud Saad El-Amari, Abdelmetalab Ferjani Tarhuni, Abdullah Ramadan Sheikhi, & Ali Ateia Elmabsout. (2021). Nutritional and health benefit knowledge of milk and dairy products consumption among medical students at Benghazi university. World Journal of Advanced Research and Reviews, 12(1), 162–174. https://doi.org/10.30574/wjarr.2021.12.1.0499

Millen, A. E., Midthune, D., Thompson, F. E., Kipnis, V., & Subar, A. F. (2006). The National Cancer Institute diet history questionnaire: validation of pyramid food servings. American journal of epidemiology, 163(3), 279–288. https://doi.org/10.1093/aje/kwj031

Milosavljević, D., Mandić, M. L., & Banjari, I. (2015). Nutritional knowledge and dietary habits survey in high school population. Collegium antropologicum, 39(1), 101-107.

Naeeni, M. M., Jafari, S., Fouladgar, M., Heidari, K., Farajzadegan, Z., Fakhri, M., ... & Omidi, R. (2014). Nutritional knowledge, practice, and dietary habits among school children and adolescents. International journal of preventive medicine, 5(Suppl 2), S171.

National Coordinating Committee on Food and Nutrition (2013), Malaysian Dietary Guidelines for Children and Adolescents 2013, Ministry of Health, Malaysia

National Coordinating Committee on Food and Nutrition (2023), Malaysian Dietary Guidelines for Children and Adolescents 2023, Ministry of Health, Malaysia

Pysz, K., Leszczynska, T., & Kopec, A. (2015). Assessment of nutritional habits and preferences of children and adolescents brought up in Krakow's orphanages. Roczniki Państwowego Zakładu Higieny, 66(3).

Racey, M., Bransfield, J., Capello, K., Field, D., Kulak, V., Machmueller, D., … & Newton, G. (2017). Barriers and facilitators to intake of dairy products in adolescent males and females with different levels of habitual intake. Global Pediatric Health, 4. https://doi.org/10.1177/2333794x17694227

Safa, T., Alvi, M. F. A., & Mahazabin, M. (2024). Nutritional Status and Lifestyle of Children in Orphanages and Middle-Class Families in Khulna City, Bangladesh: A Cross-Sectional Study. https://doi.org/10.1101/2024.07.10.24310074

Salehuddin, S., Kent, L., Ghaffar, S., Tat, C., Ibrahim, N., & Kuay, L. (2022). Milk and Milk Product Consumption: Findings From the National Health and Morbidity Survey 2022 (Adolescent Health Survey).

Sallehuddin, S. M., Talib, R. A., Salleh, R., Ambak, R., Pardi, M., Aziz, N. S. A., Wong, N. I., & Omar, M. A. (2021).

Calcium and iron intakes of adolescents in Malaysia and their relationships with body mass index (BMI): Findings from the Adolescent Nutrition Survey 2017. Malaysian Journal of Nutrition, 27(2), 317–326. https://doi.org/10.31246/MJN-2020-0054

Schiano, A. N., Nishku, S., Racette, C. M., & Drake, M. A. (2022). Parents' implicit perceptions of dairy milk and plant-based milk alternatives. Journal of dairy science, 105(6), 4946–4960. https://doi.org/10.3168/jds.2021-21626

Shaziman, S., Rani, M. D. M., Aripin, K. N. bin N., Hamid, N. A., Sulaiman, W. N. W., Rahman, Z. A., & Abdullah, M. Y. (2017). Assessing nutritional knowledge, attitudes and practices and body mass index of adolescent residents of orphanage institutions in Selangor and Malacca. Pakistan Journal of Nutrition, 16(6), 406–411. https://doi.org/10.3923/pjn.2017.406.411

Shlisky, J., Mandlik, R., Askari, S., Abrams, S., Belizan, J. M., Bourassa, M. W., Cormick, G., Driller-Colangelo, A., Gomes, F., Khadilkar, A., Owino, V., Pettifor, J. M., Rana, Z. H., Roth, D. E., & Weaver, C. (2022). Calcium deficiency worldwide: prevalence of inadequate intakes and associated health outcomes. Annals of the New York Academy of Sciences, 1512(1), 10–28. https://doi.org/10.1111/nyas.14758

Shokrvash, B., Salehi, L., Hariri Akbari, M., Ebrahimi Mamagani, M., Nedjat, S., Asghari, M., ... & Montazeri, A. (2015). Social support and dairy products intake among adolescents: a study from Iran. BMC Public Health, 15(1), 1078.

Stice, E., Marti, C. N., & Durant, S. (2011). Risk factors for onset of eating disorders: evidence of multiple risk pathways from an 8-year prospective study. Behaviour Research and Therapy, 49(10), 622-627. https://doi.org/10.1016/j.brat.2011.06.009

Tallarini, A., Zabeo, A., & Ferraretto, A. (2014). Nutritional knowledge in an Italian population of children, pre-adolescents and adolescents. Public Health Nutrition, 17(3), 708-714.

Taheri, N., Nawi, N. M., Kamarulzaman, N. H., & Abidin, Z. (2014). Factors Influencing Malaysian Purchasing Decision Of Dairy Product Based On Country Of Origin (COO). Malaysian Journal of Agricultural Economics, 27 (Dec), 39- 47

Tavares Filho, E. R., Silva, R., Pimentel, T. C., Neta, M. T. S., Prudencio, E. S., Richards, N. S., ... & Cruz, A. G. (2025). Visual and textual prompts as a nudge to promote the selection of healthy dairy products among adolescent and young adult students. International Journal of Dairy Technology, 78(3), e70043.

Wang, D., Stewart, D., Chang, C., & Shi, Y. (2015). Effect of a school-based nutrition education program on adolescents' nutrition-related knowledge, attitudes and behaviour in rural areas of China. Environmental health and preventive medicine, 20(4), 271-278.

Wham, C. A., & Worsley, A. (2003). New Zealanders' attitudes to milk: implications for public health. Public Health Nutrition, 6(1), 73-78.

ORIGINAL ARTICLE

Assessing Foodservice Quality in University Cafeterias for Health Promotion: An Importance–Performance Analysis Approach

Sharifah 'Ainur Mardhiah Syed Ahmad Shuxry[1], Noor Suzana Osman[1,2*], Wan Fathin Fariza Wan Mahmood[1,2], Ibrahim Abu Bakar [1,2],

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
Campus dining services; foodservice quality; Importance–Performance Analysis; student satisfaction; hygiene standards

ABSTRACT

Background: Campus dining services significantly influence students' dietary choices, nutritional status, and overall health. Early detection of gaps in foodservice quality allows universities to implement preventive measures that promote healthy eating behaviours and prevent nutrition-related issues. This study aimed to evaluate the importance and satisfaction of foodservice attributes at university cafeterias using the Importance–Performance Analysis (IPA) method, with a focus on identifying priority areas for preventive intervention. **Methods:** A cross-sectional survey involving 231 students was conducted across three main campus cafeterias. Data was collected using a validated questionnaire consisting of 25 attributes across food quality, food selection, food price, interactional quality, and atmospheric quality using a 7-point Likert scale. Mean scores were calculated and analyzed using the Statistical Package for the Social Sciences (IBM SPSS Statistics) Version 29. The IPA grid was used to identify priority improvement areas. **Results:** While students generally considered food and service aspects to be highly important, the findings revealed that several areas did not meet their expectations. These included food hygiene, staff friendliness, speed of service, and cleanliness of dining spaces. In contrast, food freshness, nutritional value, reasonable pricing, and staff knowledge were viewed more positively. Aspects such as entertainment and dining ambiance were less emphasised for improvement. **Conclusion:** Overall, the results indicate a gap between student expectations and the current performance of foodservice operations in key operational areas. This study highlights the need for targeted enhancements, particularly in hygiene standards. Recommended actions include comprehensive staff training, regular cleanliness inspections, and the adoption of responsive customer service protocols.

INTRODUCTION

University food services play a pivotal role in shaping students' dietary intake and contributing to their overall campus experience. The demand for nutritious and high-quality food options is increasingly pronounced within higher education institutions, as younger generations demonstrate heightened awareness of their dietary choices and actively seek healthier alternatives (Ulita et al., 2024). In response to the evolving landscape of the food industry, suppliers face mounting pressure to accommodate health-conscious consumers by offering products that balance nutritional value with taste, convenience, and variety. To effectively meet these dynamic expectations, institutions must adopt strategic approaches that prioritize customer satisfaction and align foodservice offerings with students' changing preferences.

Food service quality is critically important for university students, influencing not only their satisfaction and well-being but also their academic engagement and campus experience. Food service quality, encompassing taste, hygiene, ambiance, and presentation, has a direct impact on students' satisfaction with campus life (Hashim et al., 2024). Besides, high-quality food services ensure that students have access to nutritious meals, which are essential for maintaining energy levels, concentration, and overall health. Dissatisfaction with cafeteria services often leads students to avoid on-campus meals, potentially compromising their dietary intake (Misiran et al., 2021). Past studies found that several factors, including food quality, variety, pricing, service interactions, and the dining environment, align with student expectations, play an important role in supporting healthy eating habits and academic performance. Tan et al. (2014) in their study found that the students' satisfaction depends on the students' background, personal view, and the year of

* *Corresponding author.*
E-mail address: noorsuzana@iium.edu.my

study. Besides, factors such as age, gender, habits, and personal preferences can affect how satisfied the students are with campus food service (Shin & Yu, 2020; Tuncer et al., 2021).

According to Serhan and Serhan (2019), service quality in foodservice encompasses multiple dimensions such as food quality, food selection, pricing fairness, interactional quality, and atmospheric conditions. Each dimension contributes differently to overall satisfaction and behavioural intentions such as loyalty and repeat patronage (Andaleeb & Caskey, 2007; Tuncer et al., 2021). Therefore, students who are happy with their meals are more likely to return and tell others about the cafeteria service on campus (Pradnyadewi & Giantari, 2022).

Food hygiene and safety practices are especially critical in campus settings because of the potential for rapid transmission of foodborne illnesses in a dense population. Previous studies in university contexts have demonstrated that food quality, cleanliness, and service speed often emerge as top priorities for improvement, while factors like ambiance and seating may already meet expectations (Imran, 2018; Serhan & Serhan, 2019). However, findings vary depending on cultural, economic, and institutional factors. In Malaysia, research on campus foodservice quality revealed that food quality and service quality have significant influences on the students' satisfaction with campus foodservice (Mohd Fauzi & Rusali, 2024; Mat Nor et al., 2025). On the other hand, food price was found to be the most satisfactory attribute among international students in Malaysian research universities (Akbara et al., 2021). Food quality is widely acknowledged as a complex, relative, and multidimensional construct that encompasses nutritional, sensory, cultural, sanitary, and safety aspects, all of which contribute to consumer perceptions and expectations. Failures in areas such as hygiene, safety, or nutritional adequacy can negatively influence these perceptions and, in turn, shape the definition of quality food (Gomes et al., 2024). Thus, the Importance–Performance Analysis (IPA) method introduced by Martilla and James (1977) was used in this study to identify areas where timely action can reduce dissatisfaction and address potential food safety risk. This study addresses that gap by examining the importance and satisfaction of foodservice attributes at IIUM Kuantan and identifying key areas for improvement using the Importance–Performance Analysis (IPA) tool.

MATERIALS AND METHODS

Study Design and Data Collection

This study employed a quantitative cross-sectional design to evaluate students' satisfaction with campus foodservice and to compare perceived importance with satisfaction across service attributes. The study was conducted at the International Islamic University Malaysia (IIUM) Kuantan Campus, focusing on three main mahallah dining facilities, namely Cafeterias A, B, and C, located near student residences. Participants were recruited using convenience sampling. Inclusion criteria were local or international students (foundation, undergraduate, or postgraduate) from the IIUM Kuantan campus who purchased and consumed food at the selected cafeteria. Staff, visitors, and non-student customers were excluded. Ethical approval was obtained before data collection. Students were invited to participate either through Google Forms or paper-based questionnaires administered at the cafeterias. Informed consent was obtained from all participants.

Sample Size

The study population consists of 3,664 students across six Kulliyyah, such as Medicine, Pharmacy, Dentistry, Allied Health Sciences, Science, and Nursing. The sample size was calculated using Yamane's (1967) formula for finite populations with a 7% margin of error, as recommended for resource-constrained surveys. The minimum required sample was 193, to which an additional 20% was added to account for non-response, yielding a final target of 231 respondents.

Survey Instrument

Data were collected using a structured questionnaire adapted from Akbara et al. (2021). The instrument consisted of three sections: (a) socio-demographic information; (b) importance ratings of 25 foodservice attributes; and (c) satisfaction ratings of the same attributes. Responses were recorded on a 7-point Likert scale (1 = very unimportant/very dissatisfied to 7 = very important/very satisfied). The estimated completion time was 10–15 minutes.

Validity and Reliability of the Survey Instrument

The questionnaire was adapted and pretested to ensure content validity. The irrelevant items in the questionnaire were excluded prior to the survey. Internal consistency was assessed using Cronbach's alpha, which demonstrated excellent reliability ($\alpha = 0.93$), confirming that the

instrument was appropriate for measuring importance and satisfaction with cafeteria services.

Validity and Reliability of the Survey Instrument

The questionnaire was adapted and pretested to ensure content validity. Items that were unsuitable were removed prior to the survey. Internal consistency was assessed using Cronbach's alpha, which demonstrated excellent reliability (α = 0.93), confirming that the instrument was appropriate for measuring importance and satisfaction with cafeteria services.

Data Analysis

Data were analysed using IBM SPSS Statistics Version 29. Descriptive statistics were applied to summarise students' demographic characteristics as well as the mean ratings of importance and satisfaction for each foodservice attribute. Internal consistency of the instrument was reconfirmed using Cronbach's alpha.

The Importance–Performance Analysis (IPA) matrix was generated by combining the mean importance and satisfaction values for each cafeteria service attribute. These values were then plotted in the IPA grid to visualise priority areas for intervention (Martilla & James, 1977). The four-quadrant matrix provided an overview of attributes that required urgent improvement, those that should be maintained, areas of lesser concern, and those where resources could potentially be reallocated. Figure 1 illustrates the IPA grid and its corresponding quadrant indications.

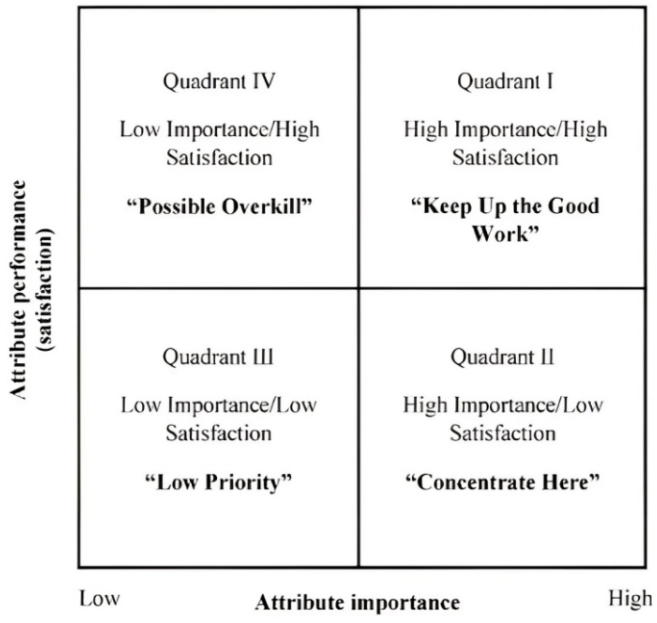


Figure 1: IPA grid for Importance- Satisfaction Analysis

RESULTS

Socio-demographic characteristics

Table 1 shows that out of 231 respondents, 227 (98.3%) were females and 4 (1.7%) were males. The majority were aged between 21–23 years. In terms of financial support, 120 (51.9%) relied on parents and 104 (45.0%) on scholarships or loans. Most respondents were from the Kulliyyah of Allied Health Sciences, with a large proportion being in their first year of study. Regarding cafeteria usage, 162 (70.1%) reported using the cafeteria daily, 53 (22.9%) used it 3–5 times per week, and 16 (6.9%) used it 1–2 times per week.

Table 1: Socio-demographic characteristics of the respondents (N=231)

Characteristics	n (%)
Age Range	
18–20	92
21–23	127
24 and above	12
Gender	
Male	227
Female	4
Sources of Expenses	
Parental support	120
Scholarship/Loan	104
Personal savings	2
Part time job income	5
Kulliyyah	
Medicine	17
Pharmacy	64
Dentistry	14
Allied Health Sciences	66
Science	23
Nursing	47
Year of Study	
1st year	96
2nd year	63
3rd year	56
4th year	15
5th and above	1
Frequency of Cafeteria Usage	
Daily	162
3–5 times per week	53
1–2 times per week	16

IPA on perceived importance and satisfaction with foodservice attributes

The IPA on the perceived importance and satisfaction with foodservice attributes is summarized in Table 2. IPA results across service quality dimensions (Interactional Quality,

Food Quality, Atmospheric Quality, Food Price, and Food Selection), showing the positioning of attributes within Quadrants I–IV, where Quadrant I indicates high importance–high performance (maintain), Quadrant II indicates high importance–low performance (improve), Quadrant III indicates low importance–low performance (low priority), and Quadrant IV indicates low importance–high performance (possible overinvestment).

Table 2: Importance-Performance Analysis for 25 Items (N=231)

Attributes	Label	Item Description	Importance	Satisfaction	Quadrant
Interactional Quality (IQ)	IQ1	Respectfulness of the foodservice outlet staff	5.64	5.34	IV
	IQ2	Staff knowledge of the food items sold in foodservice outlet	6.10	5.42	I
	IQ3	Speed of service in foodservice outlet	6.10	5.15	II
	IQ4	Cleanliness and neatness of staff appearance	6.32	5.05	II
	IQ5	Easiness talking to staff (able to answer my questions)	6.18	5.47	I
	IQ6	Friendliness of the foodservice outlet staff	6.13	5.27	II
Food Quality (FQ)	FQ1	Freshness of food ingredients	6.21	4.95	I
	FQ2	Appropriateness of food flavor	6.15	5.11	I
	FQ3	Nutritional benefits of food items	6.10	5.03	I
	FQ4	Food hygiene and safety	6.26	4.81	II
	FQ5	Appropriateness of food temperature	5.79	4.78	III
	FQ6	Attractiveness of food items displayed	5.77	4.87	III
Atmospheric Quality (AQ)	AQ1	Seat availability	5.94	6.23	I
	AQ2	Convenience of service hours	6.13	5.44	I
	AQ3	Cleanliness of the foodservice outlet	6.35	4.97	II
	AQ4	Decoration of the foodservice outlet	5.18	4.58	III
	AQ5	Noise level	5.28	4.90	III
	AQ6	Foodservice outlet entertainment, such as music	4.51	4.32	III
Food Price (FP)	FP1	The announcement of food price changes	5.94	4.81	I
	FP2	Reasonableness of the quality of food items for the price paid	6.14	5.07	I
	FP3	The attractiveness of the quantity of food items for the price paid	6.18	5.07	III
Food Selection (FS)	FS1	Variety of food served	6.08	5.13	I
	FS2	Food choices to meet nutrition needs	5.96	4.99	I
	FS3	Food choices to meet cultural preferences	5.47	5.03	IV
	FS4	Food selection to make me feel at home	5.24	4.62	III

Interactional Quality (IQ)

Within the domain of Interactional Quality (IQ), staff knowledge (IQ2) and ease of communication with staff (IQ5) were in Quadrant I (Figure 2). This indicates that customers consider these factors highly important and are satisfied with their performance, suggesting that existing standards in these areas should be maintained. Conversely, staff speed of service (IQ3), cleanliness and neatness of appearance (IQ4), and friendliness (IQ6) were positioned in Quadrant II. Since these attributes are important to customers but currently underperforming, they represent critical areas that require immediate improvement. Notably, staff friendliness (IQ6) being in Quadrant II contrasts with the ease of communication with staff (IQ5) in Quadrant I. This could imply that staff are respectful but not necessarily friendly, or that the results may reflect respondent bias. Interestingly, respectfulness

of staff (IQ1) was placed in Quadrant IV, meaning customers are satisfied with it, but do not perceive it as particularly important; thus, it requires minimal managerial attention.

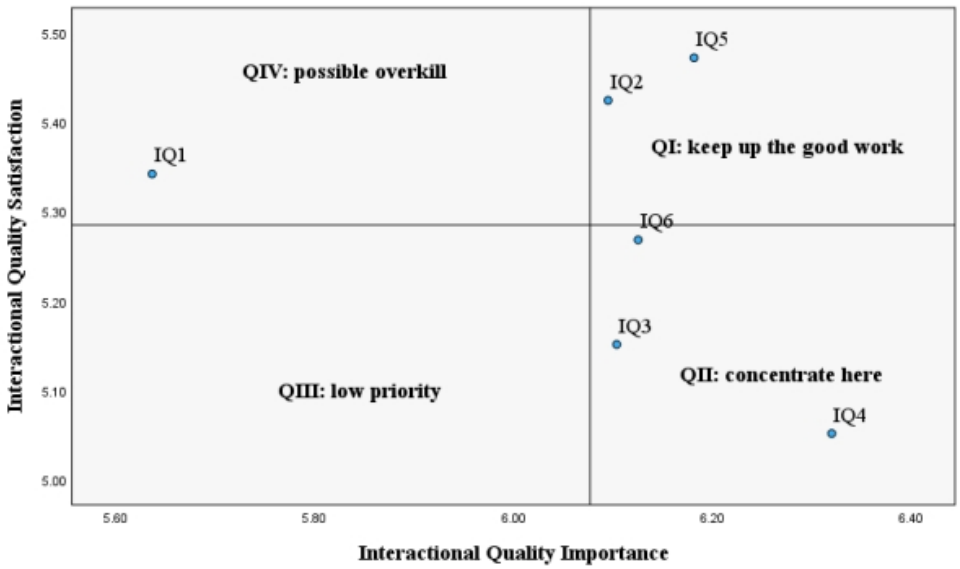


Figure 2 IPA grid for International Quality (IQ) dimensions

Food Quality (FQ)

As shown in Figure 3, the Food Quality (FQ) dimension displayed the freshness of ingredients (FQ1), suitability of food flavour (FQ2), and nutritional value of food items (FQ3) fell into Quadrant I. These are both highly important and well-performing attributes that should be sustained. In contrast, food hygiene and safety (FQ4) was mapped to Quadrant II, showing that while customers value this aspect, its current performance is unsatisfactory and demands urgent attention. Other factors, such as food temperature (FQ5) and attractiveness of food presentation (FQ6), were placed in Quadrant III, meaning they are less important and perform at lower levels, making them low-priority areas. However, considering the overall analysis, food quality as a whole was positioned in Quadrant II, implying it is a high-importance but low-satisfaction attribute and therefore warrants significant attention.

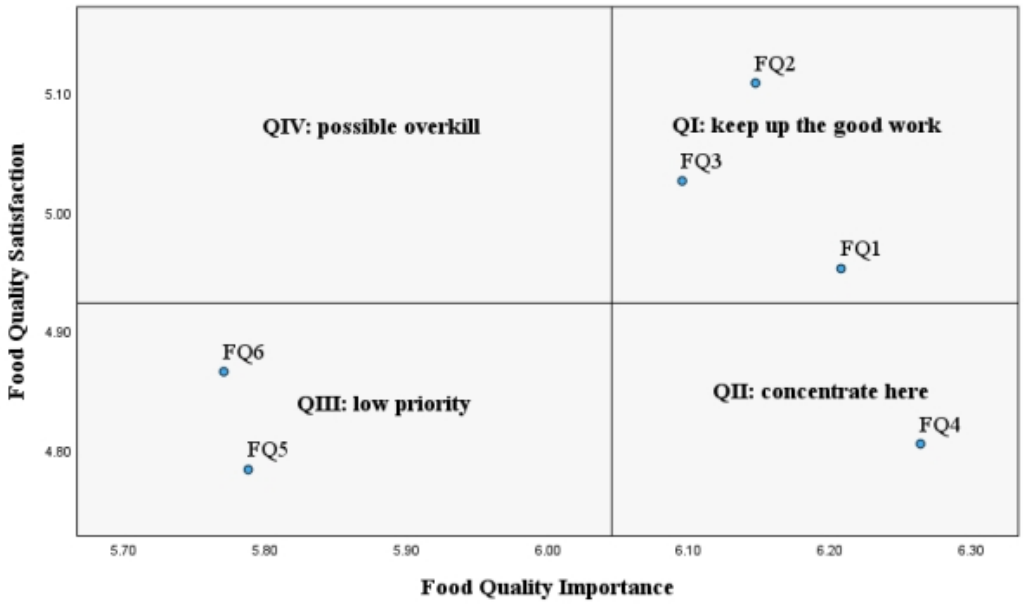


Figure 3 IPA grid for Food Quality dimensions

Atmospheric Quality

Regarding Atmospheric Quality (AQ), the overall dimension was situated in Quadrant IV, suggesting it

requires the least managerial focus (Figure 4). Nonetheless, within this category, seat availability (AQ1) and convenience of service hours (AQ2) were found in Quadrant I, indicating high importance and high satisfaction, and should therefore be preserved. On the other hand, outlet cleanliness (AQ3) was in Quadrant II, highlighting the need for urgent improvement in this essential aspect. The remaining atmospheric elements—outlet decoration (AQ4), noise level (AQ5), and entertainment features such as music (AQ6)—were positioned in Quadrant III, indicating low importance and low performance, and thus can be deprioritized.

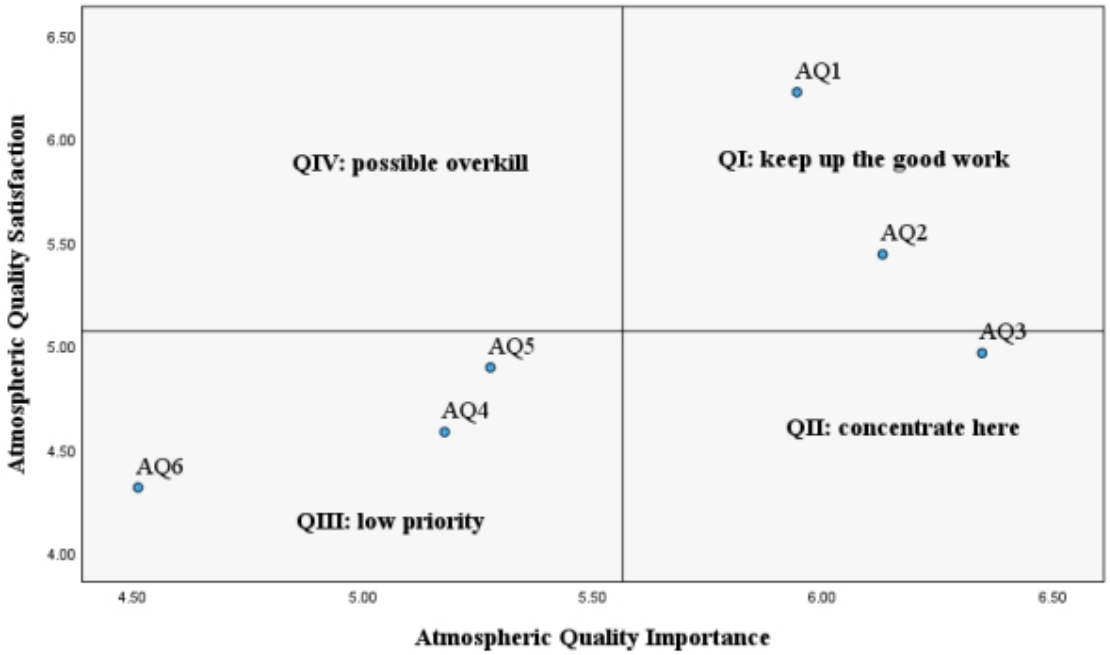


Figure 4 IPA grid for Atmospheric Quality dimensions

Food Price (FP)

With respect to Food Price (FP) as displayed in Figure 5, both fairness of food prices (FP2) and value of food quantity relative to price paid (FP3) were identified in Quadrant I, demonstrating strong performance in areas that customers deem important. In contrast, the announcement of food price changes (FP1) was classified into Quadrant III, suggesting it is of relatively low importance and does not require immediate action.

Food Selection (FS)

Finally, under Food Selection (FS), as illustrated in Figure 6, the variety of food offered (FS1) and food options meeting nutritional needs (FS2) were mapped to Quadrant I, indicating that both are important and well-delivered aspects that should be maintained. Meanwhile, food's ability to create a sense of comfort or "feeling at home" (FS4) was placed in Quadrant III, pointing to a low-priority area. Additionally, food choices accommodating cultural preferences (FS3) were found in Quadrant IV, reflecting high satisfaction for an attribute considered of low importance, suggesting little to no action is necessary at present.

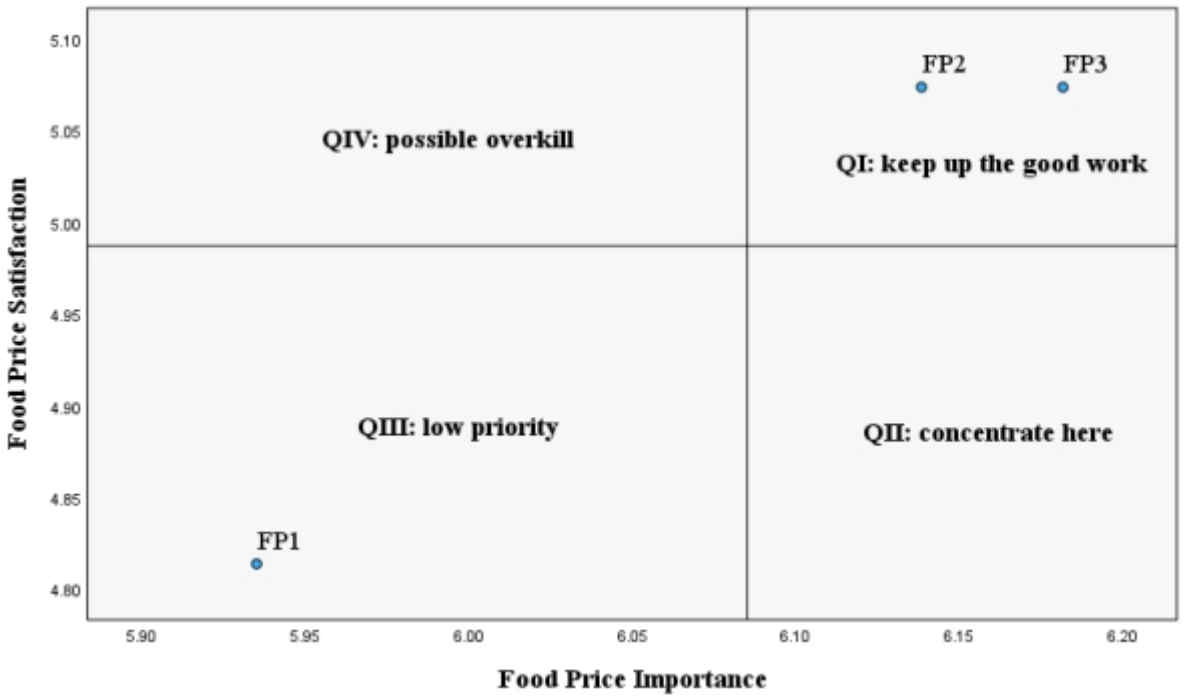


Figure 5 IPA grid for Food Price dimensions

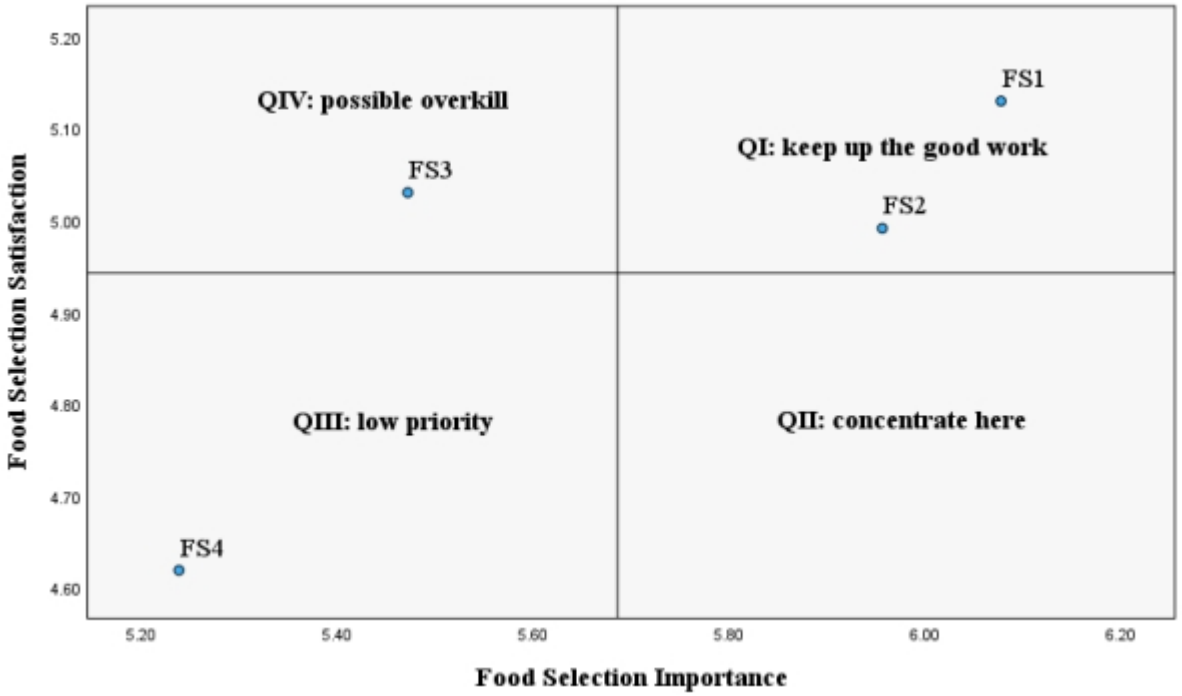


Figure 6 IPA grid for Food Selection dimensions

DISCUSSION

This study employed the Importance–Performance Analysis (IPA) to identify priority areas for enhancing foodservice quality in university cafeterias. IPA is widely used in hospitality research to compare the importance customers assign to service attributes with their satisfaction toward performance. Within the university context, this approach provides valuable managerial insight by highlighting which attributes require immediate improvement (Joung et al., 2014). The present study found that several attributes related to food quality, interactional quality, and atmospheric quality fell into the "concentrate here" quadrant, indicating that these areas were highly important to students but underperformed in practice. Interactional quality emerged as a critical determinant of

satisfaction. Attributes such as speed of service, cleanliness, and neatness of staff appearance, and ease of communication were rated as very important but scored lower in satisfaction. This result is consistent with previous studies conducted in Malaysian universities, where cleanliness, staff appearance, and efficiency were highly valued yet performed below expectations (Akbara et al., 2021). Similarly, El-Said and Fathy (2015) reported that the promptness, friendliness, and professionalism of service staff significantly influenced students' overall satisfaction in campus dining settings. These findings highlight that service interactions shape students' dining experiences and perceptions of foodservice credibility. Enhancing staff efficiency, interpersonal communication, and professionalism through systematic training and supervision is therefore essential for improving service quality in university cafeterias.

Food hygiene and safety were also identified as key determinants of satisfaction. Although students in this study viewed hygiene and safety as highly important, their satisfaction scores for these attributes were lower. This finding aligns with prior research indicating that hygiene and safety practices are central to customers' trust and satisfaction with university food services. Akbara et al. (2021) found that students expressed lower satisfaction with hygiene, nutritional value, and food quality, while El-Said and Fathy (2015) emphasized sanitation as a strong predictor of satisfaction among university diners. Similarly, Siow and Sani (2011) demonstrated that food handlers' hygiene knowledge and practices directly influence food safety outcomes in Malaysian universities. These studies collectively emphasize that ensuring strict hygiene standards, providing continuous training for food handlers, and conducting regular hygiene audits are vital to maintaining consumer confidence and safeguarding public health.

Cleanliness of foodservice outlets was another prominent attribute influencing satisfaction. A clean dining environment not only assures customers of food safety but also enhances their comfort and trust (Namkung & Jang, 2007). Conversely, poor cleanliness can result in negative perceptions and decreased patronage, even when food quality or prices are acceptable (Ha & Jang, 2010). The present study found that students rated cleanliness as one of the most important attributes, yet satisfaction remained below expectations. This is consistent with findings by Afroza and Haque (2022) and Ibrahim et al. (2018), who reported that environmental hygiene significantly affects students' evaluation of foodservice quality in Malaysian universities. Therefore, maintaining high standards of cleanliness through regular monitoring, transparent cleaning practices, and responsive maintenance should be prioritized as part of ongoing quality improvement.

Despite providing useful insights, this study has several limitations. Conducted at a single university, its findings may not be generalizable to institutions with different foodservice systems, management practices, or cultural contexts. The use of convenience sampling may have limited representativeness by excluding commuting students, staff, and cafeteria workers whose views might differ. Self-reported data could also be influenced by social desirability bias. Moreover, reliance on quantitative data restricted deeper exploration of students' experiences, while the absence of objective evaluations—such as food quality testing, pricing analysis, or hygiene audits—limited the comprehensiveness of the findings.

Future research should include a more diverse participant group encompassing commuting students, academic and administrative staff, to capture broader perspectives on campus dining experiences. Incorporating objective assessments such as food quality evaluations, price comparisons, and hygiene inspections would also enhance data validity. Longitudinal or periodic studies using consistent methodologies are recommended to track changes over time and evaluate the effectiveness of service improvements.

Overall, the IPA approach proved effective in identifying key strengths and weaknesses of university foodservice operations. The results underscore the importance of improving interactional quality, hygiene, and cleanliness to meet students' expectations. By addressing these priority areas and maintaining a culture of continuous quality improvement, universities can enhance student satisfaction, promote safer and healthier dining environments, and strengthen institutional competitiveness.

CONCLUSION

This study applied the Importance–Performance Analysis (IPA) to evaluate student satisfaction with university foodservice across five dimensions: Interactional Quality, Food Quality, Atmospheric Quality, Food Price, and Food Selection. The results showed strengths in staff knowledge, food freshness, and seating availability, but highlighted weaknesses in cleanliness, service speed, and staff friendliness. Hygiene and staff interaction require greater managerial attention, while pricing and food selection met expectations but could be more diverse and value-driven. Overall, the IPA approach effectively identified key areas for improvement to enhance student satisfaction, foodservice quality, and the university's overall competitiveness.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Afroza, K., & Haque, A. (2022). Students' satisfaction in campus cafeterias: an empirical study on public universities at Klang Valley in Malaysia. International Journal of Innovative Research and Publications, 2 (5), 58–77. https://doi.org/10.51430/IJIRP.2022.25.004

Akbara, A. Z., Chua, B. L., Han, H., & Raposo, A. (2021). Investigating international students' perception of foodservice attributes in Malaysian research universities. Sustainability (Switzerland), 13(15). https://doi.org/10.3390/su13158190

Andaleeb, S. S., & Caskey, A. (2007). Satisfaction with food services: Insights from a college cafeteria. Journal of Foodservice Business Research, 10(2), 51–65. https://doi.org/10.1300/J369v10n02_04

El-Said, O. A., Fathy, E. A. (2015). Assessing university students' satisfaction with on-campus foodservice. Tourism Management Perspectives, 16, 318–324. http://dx.doi.org/10.1016/j.tmp.2015.09.006

Gomes, T., Martinelli, S. S., Soares, P., & Cavalli, S. B. (2024). What Is Quality Food? The Opinion of Key of the Brazilian Food System. Nutrients, 16(7), 948. https://doi.org/10.3390/nu16070948

Ha, J., & Jang, S.C. (2010). Perceived values, satisfaction, and behavioral intentions: The role of familiarity in Korean restaurants. International Journal of Hospitality Management, 29(1), 2–13. https://doi.org/10.1016/j.ijhm.2009.03.009

Hashim, N. S., Othman, P. A. Z., & Jipiu, L. B. (2024). Assessing Student Satisfaction towards Food Quality and Price Fairness in UiTM Puncak Alam On-Campus Cafeteria. Journal of Tourism, Hospitality & Culinary Arts, 16(1), 514-530.
https://doi.org/10.24018/ejbmr.2022.7.2.1297

Ibrahim, M. Z. F., Mustapha, R. I. P. R., Saber, J. M., Jasni, W. N. F. W., & Mehat, N. H. (2018). A Study on Student's Satisfaction towards On-Campus Foodservice: A Case of UITM Penang Students. International Journal of Academic Research in Business and Social Sciences, 8(17), 227–237. http://dx.doi.org/10.6007/IJARBSS/v8-i17/5227

Imran, K. (2018). Investigating university students' satisfaction with on-campus cafeteria services: an empirical study in perspective of private university. Asian Journal of Empirical Research, 8(6), 225–237. https://doi.org/10.18488/journal.1007/2018.8.6/1007.6.225.237

John A. Martilla, & John C. James. (1977). Importance-Performance Analysis. Journal of Marketing.

Joung, H., Lee, D., Kim, H., & Huffman, L. M. (2014). Evaluation of the On-Campus Dining Services Using Importance-Performance Analysis. Journal of Foodservice Business Research 17(2), 136–146. https://doi:10.1080/15378020.2014.902647

Misiran, M., Md Yusof, Z., Sapirin, H., & Abdullah, I. (2022). Student Satisfaction Towards the Cafeteria in the University Campus – A Case Study. Journal of Statistical Modeling and Analytics, 4(2), 14–27.

Mohd Fauzi, N. F. ., & Rusali, R. (2024). Student Satisfaction with On-Campus Cafeteria Services at a University in Kuantan, Pahang. International Journal of Allied Health Sciences, 8(5). Retrieved from https://journals.iium.edu.my/ijahs/index.php/IJAHS/article/view/949

Namkung, Y., & Jang, S.C. (2007). Does Food Quality Really Matter in Restaurants? Its Impact On Customer Satisfaction and Behavioral Intentions. Journal of Hospitality and Tourism Research, 31(3), 387–409. http://dx.doi.org/10.1177/1096348007299924

Pradnyadewi, L. P. A., & Giantari, IG. A. K. (2022). Effect of E-Service Quality on Customer Satisfaction and Customer Loyalty on Tokopedia Customers in Denpasar. European Journal of Business and Management Research, 7(2), 200–204.

Serhan, M., & Serhan, C. (2019). The Impact of Food Service Attributes on Customer Satisfaction in a Rural University Campus Environment. International Journal of Food Science, 2019. https://doi.org/10.1155/2019/2154548

Siow, O. N., & Norrakiah Abdullah Sani. (2011). Assessment of knowledge, attitudes, and practices among food handlers at residential colleges and canteens regarding food safety. Sains Malaysiana, 40(4), 403–410.

Tuncer, I., Unusan, C., & Cobanoglu, C. (2021). Service Quality, Perceived Value and Customer Satisfaction on Behavioral Intention in Restaurants: An Integrated Structural Model. Journal of Quality Assurance in Hospitality and Tourism, 22(4), 447–475. https://doi.org/10.1080/1528008X.2020.1802390

ORIGINAL ARTICLE

Extraction of Local *Artocarpus Heterophyllus* (Jackfruit) Seeds and Its Effects on The Migration of Oral Squamous Cell Carcinoma (OSCC)

Annur Sofea Fitriah binti Rizbi[1], Mahfuzah El-Muhammady binti Abu Bakar[1] , Engku Nasiha binti Engku Ngah Sayuddin[1] , Muhammad Ashraf bin Rostam[1,2]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia.
[2]Cluster of Cancer Research Initiative IIUM (COCRII), International Islamic University Malaysia.

Keywords:
Artocarpus heterophyllus; Jackfruit seed extract; Jacalin; Oral squamous cell carcinoma; Scratch assay; Cell migration; Anticancer activity

ABSTRACT

Background: Oral squamous cell carcinoma (OSCC) is an aggressive malignancy with poor prognosis and limited therapeutic success. *Artocarpus heterophyllus* (jackfruit) is a tropical fruit rich in phytochemicals such as flavonoids, saponins, tannins, and lectins like jacalin, which have demonstrated antioxidant and anticancer properties in previous studies. Extracts from jackfruit seeds have been reported to inhibit cancer cell proliferation and metastasis through the suppression of cell migration and induction of apoptosis. This study investigates the anti-migratory effects of aqueous *Artocarpus heterophyllus* seed extract and purified jacalin on ORL-48 OSCC cell culture. **Methods**: Jackfruit seeds were first prepared and subjected to aqueous extraction to obtain the crude extract. ORL-48 cells were cultured in DMEM F-12 supplemented with FBS and antibiotics under standard incubator conditions. The cells were then either treated with the jackfruit seed crude extract, purified jacalin, cisplatin (a chemotherapy drug), or left untreated as a control. Cell migration was assessed using a scratch assay, in which a uniform wound was created on the cell monolayer. Wound closure was monitored over time to indicate cell migration. The extent of cell migration was then quantified using ImageJ software. **Results**: The ORL-48 cells that were treated with crude extract demonstrated the lowest wound closure (6.37%), followed by jacalin (13.45%), cisplatin (20.80%), and control (25.98%). Although not statistically significant ($p>0.05$), the ORL-48 cells treated with crude extract showed minimal cell migration, indicated by the least wound closure. These findings suggest the crude extract had the strongest OSCC cell inhibitory potential amongst the three treatments. **Conclusion**: The aqueous extract of jackfruit seeds exhibited potential anti-migratory properties against OSCC cells, warranting further investigation in preclinical studies as a potential anti-migratory agent for oral cancer.

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is one of the most aggressive malignancies affecting the head and neck region, accounting for a significant proportion of cancer-related morbidity and mortality worldwide (Zhou et al., 2025). Despite advances in surgical, radiotherapeutic, and chemotherapeutic interventions, the prognosis of OSCC remains poor due to its high recurrence rate, rapid metastasis, and resistance to conventional treatments. These limitations highlight the urgent need for alternative therapeutic strategies that are more effective and less toxic (Xue et al., 2025).

Natural products, particularly plant-derived bioactive compounds, have attracted considerable attention as potential anticancer agents due to their diverse pharmacological properties and relatively low side effects. *Artocarpus heterophyllus* (jackfruit), a tropical fruit widely cultivated in Malaysia, contains seeds rich in bioactive compounds, including flavonoids, saponins, and lectins such as jacalin (Fabil et al., 2024). Previous studies have demonstrated that jacalin exhibits anticancer activities in breast and colon cancer cell lines, primarily through modulation of cell proliferation, apoptosis, and migration (Geraldino et al., 2017).

However, research on the anti-migratory effects of jackfruit seed extracts and jacalin specifically on OSCC cells remains scarce. Understanding these effects is crucial, as cancer cell migration is a key step in metastasis, which significantly contributes to poor clinical outcomes. Therefore, this study aims to investigate the potential anti-migratory properties of aqueous jackfruit seed extract and purified jacalin using the ORL-48 OSCC cell line. The findings of this study may provide insights into the

[1*] *Corresponding author.*
E-mail address: ashrafrostam@iium.edu.my

therapeutic potential of jackfruit-derived compounds and support the development of novel complementary approaches for OSCC management.

MATERIALS AND METHODS

Materials

Fresh jackfruit (variety J33, Tekam Yellow) was sourced locally. The seeds were separated, washed, air-dried, and dehydrated before being ground into fine powder. Aqueous extraction was performed by soaking 10 g of powdered seeds in 100 mL distilled water for 24 hours, followed by filtration and evaporation using a rotary evaporator. The crude extract was concentrated, filtered, and stored at 4°C for use.

The OSCC cell line (ORL-48) was obtained from the Cancer Research Malaysia (CRM) and cultured in DMEM-F12 supplemented with 10% fetal bovine serum (FBS) and antibiotics.

Methods

Aqueous Extraction of Jackfruit Seeds

Jackfruit (Artocarpus heterophyllus) seeds were obtained from a local supplier in Kuantan, Pahang (Figure 1). The seeds were peeled, cleaned, and sliced into approximately 2 mm chips with the spermoderm intact. They were air-dried for three days, further dried using a food dehydrator, and then ground with a dry blender for 10 minutes to obtain a fine powder (<0.5 mm). The powder was stored in plastic pouches at room temperature until used (Mohd Ali. et al., 2015).

For extraction, 10 g of the powdered seeds were dissolved in 100 mL of distilled water (1:10 w/v) and soaked at room temperature for 24 hours. The mixture was filtered using Whatman filter paper, and the filtrate was concentrated with a rotary evaporator. The crude extract obtained was stored in an airtight container at 4 °C, then passed through a 0.45 µm syringe filter to yield a clear supernatant, which was kept at 4 °C until further analysis (Sharifah et al., 2019).

Cell Culture

ORL-48 oral squamous cell carcinoma (OSCC) cells, derived from gingival epithelium, exhibited polygonal epithelial morphology and formed cohesive monolayers. The cells were cultured in Dulbecco's Modified Eagle Medium: Nutrient Mixture F12 (DMEM F-12; Gibco, USA) supplemented with 10% fetal bovine serum (FBS) and 1% antibiotics (penicillin–streptomycin). Cultures were maintained at 37 °C in a humidified incubator with 5% CO_2 under aseptic conditions.

Frozen stocks were rapidly thawed at 37 °C, washed to remove cryoprotectant, and seeded into T25 flasks with fresh complete medium. Subculturing was performed by rinsing with phosphate-buffered saline (PBS), detaching cells with trypsin, and reseeding into new flasks. For long-term storage, cells were suspended in freezing medium (complete medium with 10% DMSO) at 1.0–1.5 × 10^6 cells/mL, aliquoted into cryovials, frozen gradually at –80 °C and transferred to liquid nitrogen for preservation.

Scratch Assay

The scratch assay was performed to assess the anti-migratory effects of crude *Artocarpus heterophyllus* seed extract and purified jacalin on ORL-48 oral squamous cell carcinoma (OSCC) cells. On Day 1, cells harvested at 80–90% confluency were seeded at 3.0 × 10^5 viable cells/well in 12-well plates with sterile glass coverslips and incubated overnight in DMEM-F12 containing 1% fetal bovine serum (FBS) to form uniform monolayers.

On Day 2, a straight wound was created using a sterile 200 µL pipette tip. Detached cells were removed by PBS washing, and the medium was replaced with serum-reduced DMEM-F12 (1% FBS). Cells were then treated with crude extract (100 mg/mL), jacalin (100 µg/mL), cisplatin (8 µg/mL), or medium alone. Treatments were administered at 1 mL per well and incubated for 24 hours at 37 °C with 5% CO_2.

On Day 3, images of the wound area were captured at 0 and 24 hours using an EVOS inverted microscope. Migration was quantified as the percentage of wound closure using ImageJ software. Each treatment was performed in triplicate, and the assay was conducted in duplicate for reproducibility.

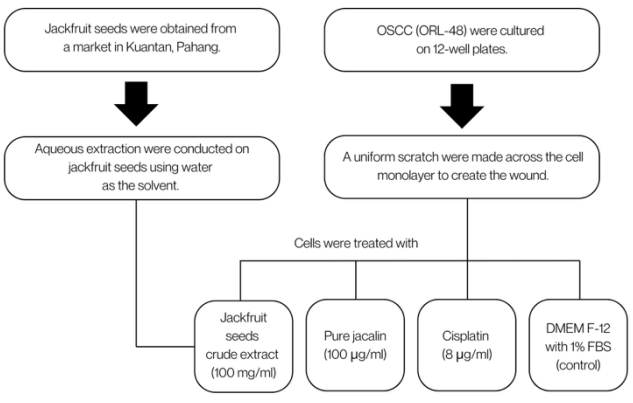


Figure 1: Research flow chart

RESULTS

The final yield of Jackfruit seed extraction was 27.5 g, corresponding to 25% of the initial seed weight, which is relatively high for aqueous extractions. The extract was concentrated to 916.7 mg/mL, providing a suitable stock solution for *in vitro* assays. Scratch assay analysis revealed that the untreated control showed the highest wound closure (25.98%), indicating normal migration. Cisplatin treatment reduced wound closure to 20.80%. Purified jacalin further decreased migration to 13.45%, while the crude extract produced the lowest closure at 6.37%, showing the strongest inhibitory trend (Table 1). Although differences were not statistically significant ($p>0.05$), the data demonstrated clear trends of reduced migration under treatment conditions.

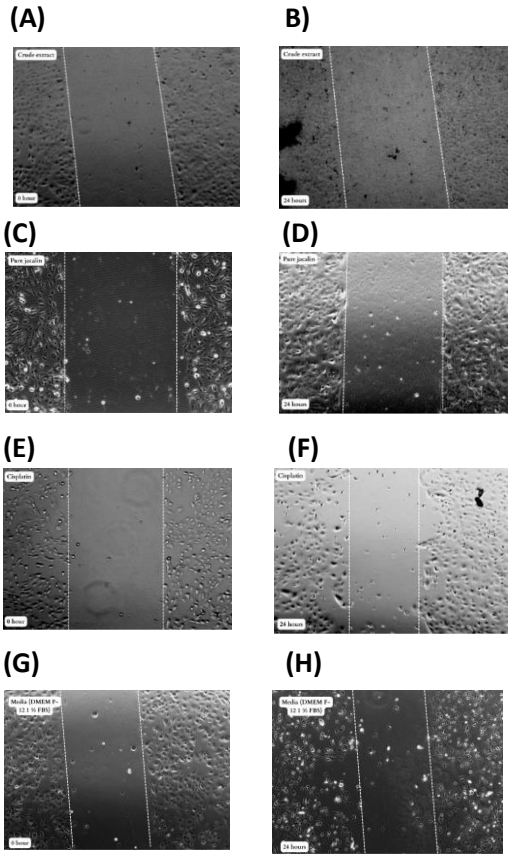


Figure 2: Wound healing assay assessing ORL-48 OSCC cell migration in response to different treatments over 24 hours. Phase-contrast microscopy images show cells treated with:

(A-B) *Artocarpus heterophyllus* aqueous extract
(C-D) Purified Jacalin
(E-F) Cisplatin
(G-H) DMEM F-12 supplemented with 1% FBS (*untreated control*)

Table 1: Summary of wound closure percentages in ORL-48 OSCC cells following different treatments. For the untreated control, cells received DMEM F-1 supplemented with 1% FBS without any exposure to cisplatin, jacalin or jackfruit crude extract.

Treatment	N	Mean (%)	Standard deviation	*Min.	Max.
Media (DMEM F-12 1% FBS)	5	25.98	22.26	2.42	57.41
Cisplatin (8 µg/ml)	6	20.80	15.43	5.36	49.60
Pure Jacalin (100 µg/ml)	6	13.45	4.41	7.14	19.19
Crude extract of jackfruit seeds (100 mg/ml)	4	6.37	2.91	2.18	8.93

**Min. = Minimum , Max.= maximum*

DISCUSSION

In this study of the ORL-48 oral squamous cell carcinoma (OSCC) line, the aqueous seed extract of *Artocarpus heterophyllus* (jackfruit) demonstrated an inhibitory trend on cell migration, while the purified lectin, Jacalin showed a weaker effect under the same conditions. The observation that the crude extract produced a stronger anti-migratory effect than the isolated jacalin suggests that multiple phytochemicals in the seed extract may act synergistically to inhibit migration more effectively than a single purified compound (Cotoraci et al., 2021).

This finding is supported by prior reviews identifying jackfruit seed bioactive compounds (lectins, flavonoids, saponins, tannins) as collectively contributing to anticancer properties (Ranasinghe et al., 2019). The superior effect of the crude extract may be explained by several possibilities. First, besides jacalin, the seed extract likely contains flavonoids, tannins and saponins that could each target different aspects of cell migration. For example, flavonoids have been shown to interfere with signalling pathways (PI3K/Akt, MAPK, NF-κB) that regulate migration and invasion (Mazuráková et al., 2024). Secondly, saponins and tannins may inhibit matrix metalloproteinases (MMPs) or reduce cell adhesion or interaction with extracellular matrix which is key steps in cell migration (Kang et al., 2008). Although our study did not assess these mechanisms directly, the phytochemical

composition of jackfruit seeds supports their plausible involvement.

Nevertheless, it is crucial to emphasise that the results did not reach statistical significance, which limits the strength of the conclusion. The possible reasons for this include the use of only a single concentration of extract or lectin, which prevents assessment of dose–response and might have missed an optimal inhibitory concentration (Cotoraci et al., 2021). Besides that, a relatively small number of replicates or lower effect size, reducing statistical power.

Other than that, the exclusive use of one cell line (ORL-48) and one migration assay format (scratch assay), which may not capture the full biological variability of OSCC. Moreover, short incubation time which is 24 hours could be one of the limitations in this study. The migration assay was conducted over a 24-hour period, which may have been too short to fully capture the anti-migratory effects of the tested compounds, especially for plant-based treatments that often act more gradually than synthetic drugs. Some bioactive compounds may require longer exposure times to interfere with migration-related signaling pathways, such as epithelial–mesenchymal transition (EMT), cytoskeletal rearrangement, or matrix metalloproteinase (MMP) expression. Without these, the observed effect remains phenotypic rather than mechanistically explained.

Based on these limitations, future work should prioritise the following, conducting dose-response experiments across multiple concentrations to determine the effective inhibitory range and calculate IC_{50}/ED_{50} values for migration suppression. Next, including time-course analysis to see if the inhibition is stable or transient over time. Incorporating mechanistic studies, for instance, assessing expression of MMP-2, MMP-9, integrins, E-cadherin/vimentin, focal adhesion kinase, or performing signalling pathway analyses (e.g., Akt, ERK phosphorylation) to uncover the underlying molecular mechanisms of migration inhibition. Moreover, testing additional OSCC cell lines (and perhaps normal oral epithelial controls) to examine whether the effect is generalisable across OSCC heterogeneity.

CONCLUSION

The aqueous extract of jackfruit seeds demonstrated the strongest anti-migratory effect on OSCC cells compared to jacalin and cisplatin. These results highlight the potential of jackfruit seeds as a source of bioactive compounds with anticancer properties. Further studies are recommended to isolate active compounds, optimize dosages, and evaluate *in viv*o efficacy for oral cancer therapy.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Higher Education, Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2022/WAB13/UIAM/02/1).

REFERENCES

Cotoraci, C., Ciceu, A., Sasu, A., Miutescu, E., & Hermenean, A. (2021). Bioactive Compounds from Herbal Medicine Targeting Multiple Myeloma. *Applied Sciences*, 11(10), 4451.

Fabil, M., Praveen Kumar Dubey, Roy, S., & Sharma, M. (2024). Jackfruit seed valorization: A comprehensive review of nutritional potential, value addition, and industrial applications. *Food & Humanity*, 3, 100406.

Geraldino, T. H., Modiano, P., Veronez, L. C., Flória-Santos, M., Garcia, S. B., & Pereira-da-Silva, G. (2017). Jacalin Has Chemopreventive Effects on Colon Cancer Development. *BioMed Research International*, 4614357.

Kang, J., Han, I., Sung, M., Yoo, H., Kim, Y., Kim, J., … & Yu, R. (2008). Soybean saponin inhibits tumor cell metastasis by modulating expressions of MMP-2, MMP-9 and TIMP-2. *Cancer Letters*, 261(1), 84-92.

Mazuráková, A., Koklesová, L., Csizmár, S. H., Samec, M., Brockmueller, A., Šudomová, M., & Golubnitschaja, O. (2024). Significance of flavonoids targeting PI3K/AKT/HIF-1α signaling pathway in therapy-resistant cancer cells – A potential contribution to the predictive, preventive, and personalized medicine. *Journal of Advanced Research*, 55, 103-118.

Mohd Ali, Z., Abdul Munaim, M. S., & Wan Sulaiman, W. M. A. (2014). Effect of lectin from *Artocarpus heterophyllus* seed on cancer cell lines. *Journal of Life Sciences & Technologies,* 2(2), 55–60.

Ranasinghe, R. A. S., Maduwanthi, S. D. T., & Marapana, R. A. U. J. (2019). Nutritional and health benefits of jackfruit (*Artocarpus heterophyllus Lam*.): A review. *International Journal of Food Science*, (1).

Sharifah F., Said, F. M., Sakinah, M., Mohamad, S., & Azizi, W. M. (2019). Backward extraction of jacalin-loaded reverse micelles into new stripping phase and toxicity evaluation on embryonic development of zebrafish.

Malaysian Journal of Biochemistry & Molecular Biology, 1(1), 134–141.

Xue, N., Wang, Y., & Wang, Z. Trends in immunotherapy for oral squamous cell carcinoma (2025). *Cell Oncology,* 48, 1159–1179.

Zhou, Y., Wang, L., & Liu, M. (2025). Oral squamous cell carcinoma: Insights into cellular heterogeneity, drug resistance, and evolutionary trajectories. *Cell Biology & Toxicol*ogy, 41, 101.

ORIGINAL ARTICLE

The Association Between Refined Carbohydrate Intake, Acne Severity, And Quality Of Life Among IIUM Students: A Cross-sectional Study

Nurin Najihah Bahrin[1], Nuraniza Azahari[1,2*], Noraishah Mohamed Nor[1,2], Aflah Afandi[1,2]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25000 Kuantan, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
Acne vulgaris; refined carbohydrates; acne; quality of life; students

ABSTRACT

Background: Acne vulgaris is a common skin condition affecting adolescents and young adults. Recent interest has grown around the role of diet, particularly refined carbohydrates in acne development. However, evidence remains mixed, and the psychosocial burden of acne is often underexplored. This study aimed to determine the pattern of refined carbohydrate intake among IIUM students, examine its association with acne severity, and assess the relationship between acne severity and quality of life. **Methods:** A cross-sectional study was conducted among IIUM undergraduate students using a self-administered questionnaire that assessed dietary intake, acne severity, and quality of life. Refined carbohydrate intake was analysed through food frequency data. Correlations between refined carbohydrate intake, acne severity, and quality of life were assessed using Pearson's correlation coefficient. **Results:** The mean daily intake of refined carbohydrates among participants was high (292.45 ± 220.35 g/day), with sugar and rice being the main contributors. However, no significant association was found between refined carbohydrate intake and acne severity. In contrast, acne severity showed a significant negative correlation with all domains of quality of life, including self-perception ($r = -0.456$), role-emotional ($r = -0.454$), role-social (r -0.369), and acne symptoms ($r = -0.602$) with $p < 0.001$. **Conclusion:** While no link was found between refined carbohydrate intake and acne severity, the results highlight the strong impact of acne on quality of life among university students. This suggests a need for more holistic acne management strategies that address both physical and psychological aspects.

INTRODUCTION

Acne vulgaris is one of the most common chronic inflammatory skin conditions, primarily affecting adolescents and young adults. Although it is not life-threatening, acne can have a significant impact on appearance, self-esteem, and psychosocial well-being. In many cases, individuals with acne experience emotional distress, low self-confidence, and social withdrawal, making it not only a dermatological concern but also a condition that affects overall quality of life (Alasbi et al., 2021; Lim et al., 2022).

The etiology of acne is multifactorial, involving genetics, hormonal regulation, environmental exposures, and lifestyle factors (Rygula et al., 2024). Among the modifiable factors, diet has received increasing attention, particularly the role of high-glycemic and refined carbohydrate foods such as rice, bread, noodles, and sugar-sweetened beverages. Refined carbohydrates are a type of carbohydrate that undergoes extensive processing, resulting in the removal of their natural fiber, vitamins, and minerals (Joshua, 2024). These foods may influence acne pathogenesis through their impact on insulin and insulin-like growth factor-1 (IGF-1) levels, leading to increased sebum production, hyperkeratinization, and inflammation (Christ et al., 2019; Melnik, 2015). However, evidence remains mixed. Some studies have reported a significant association between high refined carbohydrate intake and greater acne severity (Aslam et al., 2025; Roengritthidet et al., 2021), while others have found no clear link (Suppiah et al., 2018).

Beyond its dermatological effects, acne has been shown to adversely affect mental health and social functioning. Students with acne often report difficulties with self-perception, emotional stability, and social interactions, with more severe acne linked to poorer quality of life (Tasneem et al., 2023; Duru & Örsal, 2021). University

* *Corresponding author.*
E-mail address: nuraniza@iium.edu.my

students may be particularly vulnerable due to their transition into adulthood, exposure to academic and social pressures, and frequent consumption of carbohydrate-rich convenience foods (Pop et al., 2021; Cheng & Wong, 2021). In Malaysia, rice and sugar remain dietary staples (Ali Zainuddin et al., 2014), which may increase the relevance of exploring how refined carbohydrate intake interacts with acne severity and quality of life among students.

Despite growing research internationally, there is still limited evidence in Malaysia that simultaneously examines dietary patterns, acne severity, and their psychosocial consequences among young adults. Understanding these relationships is crucial for developing more effective, holistic strategies for acne prevention and management that address both physical and psychological aspects.

Therefore, this study aimed to identify refined carbohydrate intake among IIUM students, examine the association between refined carbohydrate intake and acne severity, and determine the relationship between acne severity and quality of life.

MATERIALS AND METHODS

This was a cross-sectional study among 377 undergraduate students from IIUM (Gombak, Kuantan, and Pagoh). The sample size was based on Krejcie and Morgan (1970), and convenience sampling was used.

Data Collection Instruments

Data were gathered through a self-administered online questionnaire consisting of four sections:
1. Sociodemographic characteristic - age, gender, kulliyyah, campus, and year of study.
2. Refined carbohydrate intake - assessed using a modified Food Frequency Questionnaire (FFQ) adapted from the Malaysian Adult Nutrition Survey 2014 (MANS), which included common local carbohydrate sources such as rice, bread, noodles, and sugar-sweetened beverages. Mean daily intake (grams/day) was calculated using standard conversion factors (Norimah, et al., 2008). 3. Acne severity - evaluated using the Global Acne Grading System (GAGS), which scores acne lesions across six anatomical regions (forehead, cheeks, nose, chin, chest, and upper back). Scores range from 0 (none) to >39 (very severe), and severity is categorised as mild, moderate, severe, or very severe (Gorelick et al., 2015).
4. Quality of life - assessed using the Acne-Specific Quality of Life (Acne-QoL) scale, which comprises four domains: self-perception, role-emotional, role-social, and acne symptoms. Items were scored on a 7-point Likert scale, with higher scores indicating less impairment (Gupta, 1998).

Ethical Approval

Ethical approval was obtained from the IIUM Research Ethics Committee (IREC 2025-KAHS/DNS17).

Statistical Analysis

Data was analyzed using SPSS Version 29. Descriptive statistics were used for sociodemographic data, refined carbohydrate intake, acne severity, and quality of life scores. Spearman's correlation test was used to examine associations between refined carbohydrate intake, acne severity, and quality of life. A p-value of <0.05 was considered statistically significant.

RESULTS

Sociodemographic Characteristics

A total of 377 students participated in this study. The majority were female (82.5%), Malay (98.4%), and aged between 21–23 years. Most participants were from the Kuantan campus (64.7%). The sociodemographic characteristics of the respondents are presented in Table 1.

Table 1: Sociodemographic Characteristics Of Respondents

Sociodemographic Variables	Frequency	n (%)
Gender		
Male	66	17.5
Female	311	82.5
Age		
18-20	96	25.5
21-23	268	71.1
>24	13	3.5
Race		
Malay	371	98.4
Indian	1	0.3
Chinese	2	05
Others	3	0.8
Campus		
Gombak	106	28.1
Kuantan	244	64.7
Pagoh	27	7.2
Year		
1	99	26.3
2	118	31.3
3	133	35.3
4	26	6.9
5	1	0.3

Refined Carbohydrate Consumption

The distribution of refined carbohydrate intake is shown in Table 2. The mean daily intake of refined carbohydrates among participants was 292.45 ± 220.35 g/day. Rice contributed the highest proportion (119.52 ± 87.63 g/day), followed by sugar (104.45 ± 156.39 g/day), bread, roti canai, noodles, pasta, and cereal.

Table 2: Refined Carbohydrate Intake among Respondents

Food items	Mean ± SD
Rice	119.52 ± 87.63
Porridge	15.36 ± 41.62
Noodle	6.12 ± 8.31
Pasta	1.20 ± 2.23
Bread	26.80 ± 58.55
Roti Canai	14.70 ± 29.08
Cereal	3.70 ± 15.13
Sugar	104.45 ± 156.39
Total CHO	292.45 ± 220.35

Acne Severity

The distribution of acne severity is shown in Table 3. Based on the Global Acne Grading System (GAGS), 80.6% of the students had mild acne, 13.5% had moderate acne, 5.6% had no acne, and only one student (0.3%) had severe acne.

Table 3: Distribution of acne severity among respondents

Food items	N (%)
None	21 (5.57)
Mild	304 (80.64)
Moderate	51 (13.53)
Severe	1 (0.26)
Very Severe	0

Quality of Life

The mean total quality of life (QoL) score was 78.44 ± 27.49. Among the four domains, acne symptoms and self-perception were the most affected. The results are presented in Table 4.

Table 4: Quality of life by domain

Domain	Mean ± SD
Self-perception	19.91 ± 8.72
Role-social	19.91 ± 8.72
Role-emotional	17.92 ± 6.73
Acne-symptoms	20.90 ± 5.63
Mean of Total Score	78.44 ± 27.49

Association Between Refined Carbohydrate and Acne Severity

The correlation analysis is presented in Table 5. There was no significant correlation between refined carbohydrate intake and acne severity (p-value < 0.05).

Table 5: Correlation and significance value between refined carbohydrate intake and acne severity

Food items	R	p-value (< 0.05)
Rice	0.016	0.753
Porridge	0.74	0.151
Noodle	0.072	0.163
Pasta	-0.005	0.916
Bread	0.082	0.111
Roti Canai	0.081	0.114
Cereal	0.016	0.751
Sugar	0.079	0.127

Association Between Acne Severity and Quality of Life

Acne severity was significantly negatively correlated with all QoL domains: self-perception (r = –0.456), role-emotional (r = –0.454), role-social (r = –0.369), and acne symptoms (r = –0.602), with all p-values < 0.001. These results are shown in Table 6.

Table 6: Correlation of acne severity and quality of life

Aspects	r	p-value (< 0.05)
Self-perception	-0.456	0.0001
Role-emotional	-0.454	0.0001
Role-social	-0.369	0.0001
Acne symptoms	-0.602	0.0001

DISCUSSION

Refined Carbohydrate Consumption and Acne Severity

The mean daily intake of refined carbohydrates among students was found to be high, with rice and sugar contributing the largest proportion. This finding is in line with the Malaysian Adult Nutrition Survey (MANS) 2014, which reported rice as the most frequently consumed staple food among Malaysian adults, with an average intake of almost 298 g/day (Ali Zainuddin et al., 2014). This reflects the cultural and dietary context in Malaysia, where rice is considered the main source of energy in daily meals.

Despite this high intake, no significant association was found between refined carbohydrate consumption and acne severity in this study. This aligns with the findings of Suppiah et al. (2018), who reported no significant associations between acne severity and the intake of sweets, fried snacks, ice cream, or sugar-sweetened beverages among Malaysian young adults. These results

suggest that refined carbohydrate intake alone may not be a strong predictor of acne development or severity, at least within this population.

However, contrasting evidence exists. A recent study by Aslam et al. (2025) demonstrated that participants with higher intakes of refined carbohydrates and sugary foods had greater acne prevalence. The proposed mechanism is that high-glycemic diets may increase insulin and insulin-like growth factor-1 (IGF-1) activity, which in turn stimulates androgen production, enhances sebum secretion, and promotes inflammation, all of which are key factors in acne pathophysiology (Melnik, 2015; Christ et al., 2019). Such discrepancies across studies may be attributed to variations in dietary assessment methods, differences in sample size and population characteristics, or genetic and hormonal factors that influence acne expression.

Although no relationship between refined carbohydrate intake and acne severity was observed in this study, the consistently high intake of refined carbohydrates remains a concern for other health outcomes. Evidence has shown that diets high in refined carbohydrates increase the risk of obesity, metabolic syndrome, insulin resistance, and type 2 diabetes (Mahzari & Mamun, 2020; Clemente-Suárez et al., 2022). Therefore, from a public health perspective, interventions promoting balanced diets with reduced refined carbohydrate consumption are still essential among university students.

Acne Severity and Quality of Life

A finding of this study was a significant negative correlation between acne severity and all quality-of-life domains, including self-perception, role-emotional, role-social, and acne symptoms. Students with more severe acne reported substantially lower quality-of-life scores, suggesting that acne exerts a broad psychosocial impact. The strongest effects were seen in the domains of self-perception and acne symptoms, highlighting the personal and visible nature of acne as a condition that directly affects confidence and self-image.

These results are consistent with previous studies across different cultural settings. Tasneem et al. (2023) reported that increasing acne severity was associated with heightened psychological distress and depressive symptoms among young adults. Similarly, Duru and Örsal (2021) and Nandy and Shrivastava (2024) emphasized the negative influence of acne on self-esteem, emotional well-being, and social participation. Together, these findings confirm that acne is not only a dermatological problem but also a condition that significantly influences mental health, social functioning, and overall well-being.

For university students, these effects may be particularly pronounced. Young adulthood is a critical period marked by academic pressures, social interactions, and identity formation. Acne during this stage may interfere with confidence in academic or professional settings, reduce social participation, and contribute to anxiety or depressive symptoms. These psychosocial burdens underline the importance of adopting a holistic approach to acne management that extends beyond pharmacological or topical treatments. Incorporating psychological counseling, peer support, or mental health services may help address the emotional and social challenges faced by students with acne.

Strengths and Limitations

This study adds to the limited Malaysian data exploring both dietary intake and quality of life in relation to acne. The use of validated tools, such as the Global Acne Grading System and Acne-QoL scale, strengthens the reliability of the findings. However, several limitations should be acknowledged. The cross-sectional design prevents causal inferences, and self-reported dietary intake may be prone to recall bias. Furthermore, the study population consisted mainly of female students from one university, which may limit generalizability to other groups.

CONCLUSION

Refined carbohydrate intake among IIUM students was high, but no significant association was found with acne severity. Acne, however, significantly reduced quality of life, affecting self-perception, emotions, social roles, and symptoms. This study highlights the importance of managing acne holistically by including psychological support, not just treating the physical condition.

ACKNOWLEDGEMENT

The authors would like to thank the Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM), for their support in facilitating this study.

REFERENCES

Alasbi, G. M., Riaz, F., Kidwan, H. N., Zanah Abdullah, Z. A., Ashfaq, T., Ghazwani, E. Y., & Ahmad, R. A. (2021). Impact of acne on quality of life among university students. *World Family Medicine Journal /Middle East Journal of Family Medicine*, *19*(2).

https://doi.org/10.5742/mewfm.2021.93985

Ali Zainuddin, A., Nadzri Jai, A., Baharuddin, A., Mahadir Naidu, B., Ying Ying, C., Siew Man, C., Nordin, F., Hashim, H., Guat Hiong, T., Yi Yi, K., Kuang Hock, L., Kuang Kuay, L., Manickam, M., Azahadi Omar, M., Fadhli Mohd Yusoff, M., Aznuddin Abd Razak, M., Fuad Mohd Anuar, M., Naim Muhammad Rasidi, M., Ani Ahmad, N., … Aris, T. (2014a):Vol. I. *National Health And Morbidity Survey 2014 : Malaysian Adult Nutrition Survey Volume I Methodology And General Findings Contributors*.

Ali Zainuddin, A., Nazri Jai, A., Husna Zainudeen, A., Baharudin, A., Mahadir Naidu, B., Ying Ying, C., Siew Man, C., Zhuo Lin, C., Paiwai, F., Nordin, F., Leng Huat, F., Jan Jan Mohamed, H. B., Akmal Abd Hamid, H., Ismail, H., Mat Hassan, H., Mutalip, H., Abu Saad, H., Tee Guat Hiong, H., Sulaiman, J., … Kaur, J. (2014). National Health and Morbidity Survey 2014b: Vol. II Survey Findings. In *National Health and Morbidity Survey*.

Aslam, M. B., Hafeez, M. A., & Khan, I. U. D. (2025). Unveiling the Role of High Glycemic Diets on Acne Formation. *Indus Journal of Bioscience Research*, *3*(3), 123–127. https://doi.org/10.70749/IJBR.V3I3.814

Cheng, S.-H., & Wong, S. E. (2021). Stress, Emotional Eating and Food Choices Among University Students During the Covid-19. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, *6*(9), 335–346. https://doi.org/10.47405/mjssh.v6i9.983

Christ, A., Lauterbach, M., & Latz, E. (2019). Western Diet and the Immune System: An Inflammatory Connection. In *Immunity* (Vol. 51, Issue 5, pp. 794–811). Cell Press. https://doi.org/10.1016/j.immuni.2019.09.020

Clemente-Suárez, V. J., Mielgo-Ayuso, J., Martín-Rodríguez, A., Ramos-Campo, D. J., Redondo-Flórez, L., & Tornero-Aguilera, J. F. (2022). The Burden of Carbohydrates in Health and Disease. In *Nutrients* (Vol. 14, Issue 18). MDPI. https://doi.org/10.3390/nu14183809

Duru, P., & Orsal, O. (2021). The effect of acne on quality of life, social appearance anxiety, and use of conventional, complementary, and alternative treatments. *Complementary Therapies in Medicine*, *56*. https://doi.org/10.1016/j.ctim.2020.102614

Gorelick, J., Daniels, S. R., Kawata, A. K., Degboe, A., Wilcox, T. K., Burk, C. T., & Douse-Dean, T. (2015). Acne-Related Quality of Life Among Female Adults of Different Races/Ethnicities. *Journal of the Dermatology Nurses' Association*, *7*(3), 154–162. https://doi.org/10.1097/JDN.0000000000000129

Gupta, M. A., Johnson, A. M., & Gupta, A. K. (1998). The development of an Acne Quality of Life scale: reliability, validity, and relation to clinical severity. British Journal of Dermatology, 139(3), 435–440.

Joshua, J. (2024). Deciphering carbohydrate quality: A guide to making healthier dietary choices. *Journal of Experimental Food Chemistry, 10*(02), 473. https://doi.org/10.37421/2472-0542.2024.10.473

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement, 30*(3), 607–610.

Lim T H, Badaruddin N S F, Foo S Y, M A Bujang, & P Muniandy. (2022). *Prevalence and psychosocial impact of acne vulgaris among high school and university students in Sarawak, Malaysia*.

Mahzari, M., & Mamun, A. (2020). Does consumption of refined carbohydrates predict the incidence of type 2 diabetes mellitus? a systematic review and meta-analysis. In *Romanian Journal of Diabetes, Nutrition and Metabolic Diseases* (Vol. 27, Issue 2, pp. 168–179). Romanian Society of Diabetes, Nutrition and Metabolic Diseases. https://doi.org/10.46389/rjd-2020-1027

Melnik, B. C. (2015). Linking diet to acne metabolomics, inflammation, and comedogenesis: An update. In *Clinical, Cosmetic and Investigational Dermatology* (Vol. 8, pp. 371–388). Dove Medical Press Ltd. https://doi.org/10.2147/CCID.S69135

Nandy, P., & Shrivastava, T. (2024). Exploring the Multifaceted Impact of Acne on Quality of Life and Well-Being. *Cureus*. https://doi.org/10.7759/cureus.52727

Norimah AK, Safiah M, Jamal K, Haslinda S, Zuhaida H, Rohida S, et al. Food consumption patterns: findings from the Malaysian Adult Nutrition Survey (MANS). Mal J Nutr. 2008;14:25–39.

Pop, L. M., Iorga, M., Muraru, I. D., & Petrariu, F. D. (2021). Assessment of Dietary Habits, Physical Activity and Lifestyle in Medical University Students. *Sustainability 2021, Vol. 13, Page 3572*, *13*(6), 3572. https://doi.org/10.3390/SU13063572

Roengritthidet, K., Kamanamool, N., Udompataikul, M., Rojhirunsakool, S., Khunkhet, S., & Kanokrungsee, S. (2021). Association Between Diet and Acne Severity: A Cross-sectional Study in Thai Adolescents and Adults. *Acta Dermato-Venereologica*, *101*(12). https://doi.org/10.2340/actadv.v101.569

Ryguła, I., Pikiewicz, W., & Kaminiów, K. (2024). Impact of Diet and Nutrition in Patients with Acne Vulgaris. In *Nutrients* (Vol. 16, Issue 10). https://doi.org/10.3390/nu16101476

Suppiah, T. S. S., Sundram, T. K. M., Tan, E. S. S., Lee, C. K., Bustami, N. A., & Tan, C. K. (2018a). Acne vulgaris and its association with dietary intake: A Malaysian perspective. *Asia Pacific Journal of Clinical Nutrition*, *27*(5), 1141–1145. https://doi.org/10.6133/apjcn.072018.01

Tasneem, T., Begum, A., Chowdhury, M. R. K., Rahman, S., Macassa, G., Manzoor, J., & Rashid, M. (2023). Effects

of acne severity and acne-related quality of life on depressive symptoms among adolescents and young adults: a cross-sectional study in Bangladesh. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1153101

ORIGINAL ARTICLE (ORIGINAL ARTICLE)

Nutritional and Functional Characterization of Malaysian Banana Flours (Pisang Berangan and Pisang Nangka) and Their Application in High-Fiber Biscuit Development

Normah Haron[1], Hanapi Mat Jusoh[2], Noraslinda Muhamad Bunnori [1*]

[1] Department of Biotechnology, Kulliyyah of Science, International Islamic University Malaysia (IIUM), Jalan Sultan Ahmad Shah, 25200 Kuantan, Pahang, Malaysia.
[2] Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM), Jalan Sultan Ahmad Shah, 25200 Kuantan, Pahang, Malaysia.

ABSTRACT

Background: Banana flour from unripe bananas is a promising functional ingredient due to its distinctive nutritional composition and bioactive compounds. This study evaluated the nutritional and functional potential of banana flour from two Malaysian varieties, *Musa acuminata* (Pisang Berangan) and *Musa paradisiaca* (Pisang Nangka), with emphasis on their use in value-added bakery products. **Methods:** Unripe bananas were processed into flour by freeze-drying, followed by proximate and functional property analyses.Biscuits were formulated using Pisang Berangan flour (100%, 50% banana–wheat, and 50% banana–rice blends) for sensory evaluation, while both Pisang Berangan and Pisang Nangka flours were analyzed for nutritional composition and physicochemical properties. A sensory evaluation was conducted among 30 randomly selected undergraduate students from IIUM Kuantan to assess three biscuit formulations (banana flour, banana–wheat flour, and banana–rice flour) for aroma, texture, color, flavor, and overall acceptability using a 9-point hedonic scale (1 = dislike extremely; 9 = like extremely). Samples were coded with two-digit numbers, and participants ranked their preferences and provided additional feedback. **Results:** Pisang Berangan flour contained significantly higher ($p < 0.05$) protein (3.74%), crude fiber (5.22%), and ash (2.55%) than Pisang Nangka (2.89%, 3.18%, and 2.02%, respectively). Carbohydrate was the major fraction in both flours (>90%), with Pisang Berangan exhibiting higher starch (92.06%) and amylose (67.71%) compared to Pisang Nangka (88.74% and 59.83%, respectively), indicating stronger thickening and stabilizing potential. Sensory evaluation showed that the 50% banana flour–wheat flour formulation obtained the highest scores for flavor (6.93 ± 1.53), texture (6.87 ± 1.25), and overall acceptability (7.00 ± 1.53). In contrast, the 100% Pisang Berangan flour biscuits exhibited superior nutritional quality but were less preferred in terms of texture. **Conclusion:** Overall, a 50% replacement of wheat flour with Pisang Berangan flour provides the best compromise between texture, flavor, and nutrient density, highlighting its potential as a sustainable functional ingredient for the development of high-fiber bakery products.

Keywords:
Banana flour; functional foods; resistant starch; bioactive compounds; product development; sustainability

INTRODUCTION

Bananas (*Musa* spp.) are among the world's most important fruit crops, with global production exceeding 125 million tonnes annually, primarily in tropical and subtropical regions (FAO, 2023). In Malaysia, bananas are the second most widely cultivated fruit, yet their high perishability leads to substantial post-harvest losses estimated at 20–30% (FAO, 2023). Converting bananas into flour has been proposed as an effective strategy to extend shelf life while preserving nutritional quality (Bezerra et al., 2013).

Banana flour is rich in resistant starch, dietary fiber, and bioactive compounds, which have been associated with improved glycemic control, gut health, and satiety (Menezes et al., 2020; Munir et al., 2024). Moreover, it is naturally gluten-free, making it suitable for individuals with celiac disease or gluten intolerance (Shiga et al., 2020). Previous studies have highlighted its potential in baked goods, pasta, and snack formulations (Agama-Acevedo et al., 2009; Biernacka et al., 2020; Ovando-Martínez et al., 2009). However, varietal differences in nutrient composition and functional properties can significantly influence flour performance in food applications (Anyasi et al., 2019; Bi et al., 2017).

[*] *Corresponding author.*
E-mail address: hanapi@iium.edu.my; noraslinda@iium.edu.my

In Malaysia, *Musa acuminata* (Pisang Berangan) and *Musa paradisiaca* (Pisang Nangka) are widely cultivated varieties, yet their suitability for flour production has received limited scientific attention. While banana flour has been investigated globally, the nutritional and functional potential of these Malaysian varieties remains poorly characterized, particularly in relation to their application in baked products such as biscuits.

Therefore, the objective of this study was to evaluate the nutritional and functional properties of Pisang Berangan and Pisang Nangka flours and to assess their application in high-fiber biscuit development.

MATERIALS AND METHODS

Preparation of Banana Flour

Fresh unripe green bananas (Pisang Berangan and Pisang Nangka) were thoroughly washed with tap water to remove adhering dust and soil. The fruits were peeled, cut into approximately 1 cm slices, and immediately immersed in a solution containing 1% (w/v) ascorbic acid and 0.185% (w/v) EDTA to prevent enzymatic browning. The slices were frozen at −80 °C for 6 days, followed by freeze-drying for 6 days. The dried samples were ground into fine powder using an electric grinder and sieved through a 250 µm mesh. The flour was stored in sealed plastic containers at 25 °C until further analysis.

For the biscuit formulation, a total of 3.66 kg of fresh unripe bananas (*Musa paradisiaca*, variety Pisang Berangan) at the mature green stage were procured from the Kuantan market for flour preparation. The remaining baking ingredients were sourced from a local supermarket.

Formulation of Biscuits

The biscuit formulation was adapted from Chavan et al. (2020) with modifications in flour ratios. Margarine (50 g) and powdered sugar (60 g) were creamed with a spatula until smooth, followed by the addition of one egg and one teaspoon of vanilla extract. Depending on formulation, banana flour, banana–wheat flour (50:50), or banana–rice flour (50:50) was incorporated gradually to form a uniform dough. The dough was rolled to 0.5 cm thickness, cut into uniform shapes, and baked at 180 °C for 15 minutes. Biscuits were cooled at room temperature, packed in airtight containers, and stored for further analyses.

Proximate Analysis of Banana Flours

Proximate analyses were carried out independently for each type of banana flour, namely Pisang Berangan and Pisang Nangka. Proximate composition was determined according to AOAC Official Methods (AOAC, 2016). Protein was analyzed by the Kjeldahl method, moisture content by oven drying at 105 °C, and ash content by incineration in a muffle furnace at 550 °C. Lipids were extracted using Soxhlet with hexane, and crude fiber was quantified by sequential acid and alkali digestion. Carbohydrate content was estimated using the phenol–sulphuric acid colorimetric method. Starch content was determined with the Rapid Total Starch Assay Kit (Megazyme), while amylose was measured gravimetrically (Subroto et al., 2020), with amylopectin calculated by difference.

Functional Properties of Banana Flour

Functional properties were assessed to characterize flour performance. Starch clarity was measured as percentage transmittance at 650 nm after gelatinization and cooling (Achille et al., 2007). Viscosity and pasting behavior were determined using a vibro-viscometer (SV-10, AND Co. Ltd.) during heating and cooling cycles (Khoozani et al., 2019).

Sensory Analysis of Banana Flour (Pisang Berangan) Biscuits

Thirty undergraduate students at IIUM Kuantan participated in the sensory evaluation. Participants provided informed consent and were supplied with water to cleanse their palates between samples. Three biscuit formulations (100% banana flour, banana–wheat flour, and banana–rice flour) were evaluated for aroma, texture, color, flavor, and overall acceptability using a nine-point hedonic scale (1 = dislike extremely, 9 = like extremely). Samples were coded with two-digit random numbers. Participants also ranked their preferences and provided open comments.

Statistical Analysis

All experiments were performed in triplicate unless otherwise stated. Data were reported as mean ± standard deviation (SD). Independent t-tests were used to compare flour varieties, while one-way ANOVA with Tukey's HSD post hoc test was applied to compare biscuit formulations. Sensory data were analyzed using repeated-measures ANOVA with Bonferroni adjustments or Friedman's test when assumptions were not met. Analyses were conducted in SPSS

Version 27 (IBM Corp., Armonk, NY), with significance set at $p < 0.05$.

RESULTS

Nutritional Composition of Banana Flours

The proximate composition of banana flours from Pisang Berangan and Pisang Nangka is presented in Table 1. Both varieties showed low moisture content (<1.5%), indicating stability for storage. Pisang Berangan exhibited significantly higher ($p < 0.05$) protein (3.74%), crude fiber (5.22%), and ash (2.55%) compared with Pisang Nangka (2.89%, 3.18%, and 2.02%, respectively). Carbohydrate was the predominant fraction (>90%) in both, with Pisang Berangan also showing significantly higher starch (92.06%) and amylose (67.71%).

Table 1: Proximate composition of banana flours (mean ± SD, n = 3)

Parameter (%)	Pisang Berangan	Pisang Nangka	p-value
Moisture	1.33 ± 0.05	1.41 ± 0.07	0.218
Ash	2.55 ± 0.11	2.02 ± 0.08	0.034*
Protein	3.74 ± 0.12	2.89 ± 0.09	0.021*
Crude fiber	5.22 ± 0.15	3.18 ± 0.13	0.012*
Crude fat	0.97 ± 0.03	0.85 ± 0.04	0.147
Carbohydrate (diff.)	91.84 ± 0.27	93.07 ± 0.31	0.086
Starch	92.06 ± 1.12	88.74 ± 1.25	0.041*
Amylose	67.71 ± 1.04	59.83 ± 1.11	0.029*

*Note. Values are mean ± standard deviation (SD). *$p < 0.05$ indicates significant difference between varieties (independent t-test).*

Functional Properties of Banana Flours

Functional evaluations revealed that Pisang Berangan flour exhibited significantly higher viscosity, gelatinization temperature, water absorption capacity, and pasting clarity than Pisang Nangka ($p < 0.05$). These results highlight the superior functional potential of Pisang Berangan for food applications.

Table 2: Functional properties of banana flours (mean ± SD, n = 3)

Property	Pisang Berangan	Pisang Nangka	p-value
Viscosity (mPa·s)	10.92 ± 0.34	5.54 ± 0.29	0.008*
Gelatinization Temp (°C)	78.3 ± 0.90	75.6 ± 0.80	0.041*
Water absorption (g/g)	2.87 ± 0.12	2.31 ± 0.14	0.019*
Pasting clarity (Abs620)	0.42 ± 0.01	0.35 ± 0.02	0.027*

*Note. Values are mean ± SD. *$p < 0.05$ indicates significant difference (independent t-test).*

Nutritional Composition of Pisang Berangan Flour–Based Biscuit Formulations

The proximate composition of biscuits varied depending on flour formulation (Table 3). Biscuit A (100% banana flour) had the highest crude fiber (4.96%) and carbohydrate content (88.50%). Biscuit B (50% banana + wheat flour) had significantly higher protein (4.32%) but lower carbohydrate (86.30%). Biscuit C (50% banana + rice flour) showed intermediate values.

Table 3: Nutritional composition of biscuit formulations (mean ± SD, n = 3)

Parameter (%)	Biscuit A (100% Banana)	Biscuit B (50% Banana + Wheat)	Biscuit C (50% Banana + Rice)	p-value
Moisture	2.14 ± 0.06^{b}	2.48 ± 0.08^{a}	2.37 ± 0.07^{ab}	0.032*
Ash	1.94 ± 0.09^{a}	1.76 ± 0.07^{b}	1.81 ± 0.06^{b}	0.041*
Protein	3.21 ± 0.11^{c}	4.32 ± 0.13^{a}	2.89 ± 0.10^{b}	0.019*
Crude fiber	4.96 ± 0.15^{a}	3.11 ± 0.14^{c}	3.54 ± 0.13^{b}	0.015*
Carbohydrate	88.50 ± 0.27^{a}	86.30 ± 0.31^{c}	87.40 ± 0.29^{b}	0.027*

Note. Superscripts indicate significant differences among formulations (Tukey's post hoc test, $p < 0.05$).

Sensory Evaluation of Biscuits

Aroma and color did not differ significantly among the formulations. However, significant differences were

observed in texture, flavor, and overall acceptability ($p < 0.05$). Biscuits formulated with banana and wheat flour (Biscuit B) consistently achieved the highest scores, while 100% banana flour biscuits scored lowest in texture and acceptability.

Table 4: Sensory evaluation scores of biscuits formulated with banana flour and blends (mean ± SD, n = 30)

Attribute	Biscuit A (100% Banana)	Biscuit B (Banana + Wheat)	Biscuit C (Banana + Rice)	p-value
Aroma	6.37 ± 1.79	6.80 ± 1.49	6.40 ± 1.61	0.524
Texture	5.30 ± 1.90	6.87 ± 1.25*	5.90 ± 1.79	0.002*
Color	7.20 ± 1.35	7.23 ± 1.00	6.60 ± 1.30	0.086
Flavor	4.53 ± 2.31	6.93 ± 1.53*	6.30 ± 1.96	<0.001*
Overall acceptability	5.13 ± 2.08	7.00 ± 1.53*	6.23 ± 1.65	<0.001*

*Note. *$p < 0.05$ indicates significant differences (one-way ANOVA).*

DISCUSSION

Nutritional Composition of Banana Flours

The nutritional profile of Pisang Berangan flour was superior to Pisang Nangka, with significantly higher protein, crude fiber, and ash contents. Dietary fiber is strongly associated with reduced risk of non-communicable diseases such as type 2 diabetes, obesity, and cardiovascular disorders (Falcomer et al., 2019). The higher starch and amylose contents in Pisang Berangan flour align with earlier studies showing that unripe banana flour is a rich source of resistant starch (Ovando-Martínez et al., 2009; Khoozani et al., 2019). Resistant starch provides functional health benefits, including glycemic regulation, satiety promotion, and colonic fermentation (Menezes et al., 2020; Munir et al., 2024).

Functional Properties and Food Processing Implications

Functional property analyses further confirmed the superior performance of Pisang Berangan flour compared to Pisang Nangka. Its significantly higher viscosity, gelatinization temperature, and water absorption capacity enhance texture, moisture retention, and stability in baked goods and other processed food applications (Agama-Acevedo & Bello-Pérez, 2017). The higher amylose content contributes to retrogradation stability, which is advantageous for shelf-life extension. This is consistent with Sarawong et al. (2014), who found that extrusion of amylose-rich banana flour improved functional properties without compromising antioxidant capacity.

Nutritional Enhancement of Biscuit Formulations

Banana flour incorporation into biscuits improved nutritional value, particularly fiber enrichment in the 100% banana flour biscuits. However, complete substitution negatively affected texture and overall acceptability, consistent with findings by Dudu et al. (2021) and Khoozani et al. (2020). Partial substitution with wheat flour (50%) yielded the most balanced outcome, with improved protein content, acceptable texture, and favorable sensory scores. This aligns with Shiga et al. (2020), who reported that blending banana flour with wheat improved consumer acceptance compared to full replacement.

Sensory Perception and Consumer Acceptance

Texture was the most decisive factor influencing biscuit acceptability, with the banana–wheat flour formulation outperforming the others. The blending strategy retained the desirable lightness of wheat flour while providing functional enrichment from banana flour. This supports the conclusion that flour blends allow producers to improve nutrition without compromising sensory quality (Anyasi et al., 2019). Moreover, the distinctive flavor contribution of banana flour was generally well accepted, highlighting its market potential in bakery products.

Limitations of the Study

A limitation of the present study is the relatively small sensory panel size (n = 30), which consisted of undergraduate students. While trained panels ensure consistency, larger consumer panels representing diverse demographics would be more reflective of market acceptance. Therefore, the sensory results should be interpreted with caution, and broader consumer testing is warranted.

Future Research Directions

Further research should explore technological interventions to improve 100% banana flour products. Strategies such as hydrocolloid addition (xanthan gum, guar gum, carrageenan), enzymatic modification, or pre-gelatinization could improve texture and

palatability. Blending banana flour with protein-rich flours such as soy or chickpea may also enhance nutritional and sensory attributes. In addition, consumer-scale acceptance studies and biochemical profiling of bioactive compounds are needed to validate the functional benefits of banana flour.

Sustainability and Food Security Perspective

Banana flour production contributes to reducing agricultural losses and enhancing food security. Bananas are among the most wasted fruits globally due to perishability, with post-harvest losses estimated at 20–30% (FAO, 2023). Converting unripe bananas into flour extends shelf life, supports smallholder farmers, and aligns with the United Nations Sustainable Development Goal 12 on responsible consumption and production (Viana et al., 2024). In Malaysia, banana flour innovation also reduces reliance on imported wheat flour, supporting local food industries and national food security initiatives.

CONCLUSION

This study demonstrated that banana flour, particularly from the Pisang Berangan variety, possesses superior nutritional and functional qualities compared to Pisang Nangka, characterized by higher protein, fiber, and amylose contents. When incorporated into biscuit formulations, Pisang Berangan flour significantly enhanced the nutritional profile, with the 50% banana–wheat flour blend yielding the most favorable balance between sensory acceptability and nutrient enrichment. Complete substitution with banana flour improved fiber content but reduced textural preference. These findings underscore the potential of Pisang Berangan flour as a sustainable and functional ingredient for developing high-fiber bakery products, contributing to waste reduction, local food innovation, and improved food security in Malaysia.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Agama-Acevedo, E., Islas-Hernández, J. J., Pacheco-Vargas, G., Osorio-Díaz, P., & Bello-Pérez, L. A. (2009). Pasta with unripe banana flour: Physical, texture, and preference study. *Journal of Food Science, 74*(6), S263–S267. https://doi.org/10.1111/j.1750-3841.2009.01215.x

Anyasi, T. A., Jideani, A. I. O., & Mchau, G. R. A. (2019). Effect of organic acid pretreatment on functional properties of unripe banana flour and optimized snack formulations. *Food Bioscience, 28*, 63–74. https://doi.org/10.1016/j.fbio.2019.01.003

Bezerra, C. V., Rodrigues, A. M. C., Amante, E. R., & da Silva, L. H. M. (2013). Nutritional potential of green banana flour obtained by drying in spouted bed. *Revista Brasileira de Fruticultura, 35*(4), 1140–1148. https://doi.org/10.1590/S0100-29452013000400025

Biernacka, B., Dziki, D., Różyło, R., & Gawlik-Dziki, U. (2020). Banana powder as an additive to common wheat pasta. *Foods, 9*(1), 53. https://doi.org/10.3390/foods9010053

Dudu, O. E., Idowu-Adebayo, F., & Akinola, A. A. (2021). Nutritional, functional and sensory properties of cookies produced from composite flours of banana, wheat and soy. *Food Research, 5*(5), 182–189. https://doi.org/10.26656/fr.2017.5(5).413

Falcomer, A. L., de Lima, B. R., Ginani, V. C., Zandonadi, R. P., & Botelho, R. B. A. (2019). Health benefits of green banana flour: A systematic review. *Nutrients, 11*(6), 1222. https://doi.org/10.3390/nu11061222

FAO. (2023). *Banana market review 2023*. Food and Agriculture Organization of the United Nations. https://www.fao.org/3/myz100en/myz100en.pdf

Khoozani, A. A., Bekhit, A. E.-D. A., & Birch, J. (2019). Effects of different drying conditions on the starch content, thermal properties and some physiochemical parameters of whole green banana flour. *International Journal of Biological Macromolecules, 130*, 938–946. https://doi.org/10.1016/j.ijbiomac.2019.03.010

Khoozani, A. A., Kebede, B., Birch, J., & Bekhit, A. E.-D. A. (2020). The effect of bread fortification with whole green banana flour on its physicochemical, nutritional and in vitro digestibility. *Foods, 9*(2), 152. https://doi.org/10.3390/foods9020152

Menezes, E. W., Dan, M. C. T., Cardenette, G. H. L., & Lajolo, F. M. (2020). Resistant starch in foods: Current advances and challenges. *Current Opinion in Food Science, 33*, 118–124. https://doi.org/10.1016/j.cofs.2020.01.007

Munir, H., Begum, R., Farooq, S., & Gillani, S. (2024). Green banana resistant starch: A promising potential

as functional ingredient against certain maladies. *Food Science & Nutrition, 12*, e3787. https://doi.org/10.1002/fsn3.3787

Ovando-Martínez, M., Sáyago-Ayerdi, S. G., Agama-Acevedo, E., Goñi, I., & Bello-Pérez, L. A. (2009). Unripe banana flour as an ingredient to increase the undigestible carbohydrates of pasta. *Food Chemistry, 113*(1), 121–126. https://doi.org/10.1016/j.foodchem.2008.07.035

Sarawong, C., Schoenlechner, R., Sekiguchi, K., Berghofer, E., & Ng, P. K. (2014). Effect of extrusion cooking on the physicochemical properties, resistant starch, phenolic content and antioxidant capacities of green banana flour. *Food Chemistry, 143*, 33–39. https://doi.org/10.1016/j.foodchem.2013.07.081

Shiga, T. M., Lajolo, F. M., & Menezes, E. W. (2020). Development of gluten-free bakery products using unripe banana flour: Nutritional and technological insights. *Journal of Cereal Science, 96*, 103126. https://doi.org/10.1016/j.jcs.2020.103126

Viana, L. M., de Souza, T. S. P., Silva, L. M. R., et al. (2024). Green banana (Musa ssp.) mixed pulp-and-peel flour: A strategy to reduce post-harvest losses and develop healthy foods. *Food Chemistry*. Advance online publication. https://doi.org/10.1016/j.foodchem.2024.138015

ORIGINAL ARTICLE

Food Neophobia Among Preschool Children in Kuantan, Pahang

Nurul Syamin Mohamed Kamal[1], Pramitha Sari[2], Nurul Hazirah Jaafar[1,3,*]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2]Department of Nutrition, Faculty of Health Sciences, Universitas Alma Ata, Yogyakarta, Indonesia
[3]Children Health and Wellbeing Research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
food neophobia; preschool children; food consumption; association

ABSTRACT

Background: Food neophobia, the reluctance to try unfamiliar foods, is common in young children and has been linked to unhealthy dietary habits, including low fiber intake and preference for energy-dense foods. In Malaysia, evidence on food neophobia among preschool-aged children remains limited, despite persistent concerns about child malnutrition, particularly stunting. This study aimed to assess food neophobia, dietary patterns, and their associations with sociodemographic and nutritional factors among preschool children in Kuantan, Pahang. **Methods:** A cross-sectional study was conducted among 44 children aged 4–6 years and their caregivers. Data were obtained through caregiver-administered, face-to-face questionnaires comprising the Child Food Neophobia Scale (CFNS), a Food Frequency Questionnaire (FFQ), and sociodemographic information. Children's anthropometric measurements were collected to assess their nutritional status. Data analysis was performed using descriptive statistics and chi-square tests with SPSS version 20**. Results:** In this study, children were categorized into low (25.0%), moderate (52.3%), and high (22.7%) food neophobia. The highest dietary intake was for beverages and cereal products, while the lowest was for nuts and bread spreads. Their nutritional status showed 6.8% underweight, 22.7% stunted, and 6.8% wasted. No significant associations were found between food neophobia and nutritional status or most sociodemographic variables. **Conclusion:** Although no significant associations were found, the presence of stunting and wasting underscores the importance of promoting dietary variety from early childhood. Early dietary interventions and active caregiver engagement are essential to improve food acceptance and support optimal growth and health among preschool-aged children.

INTRODUCTION

Food neophobia, commonly observed in early childhood, refers to the reluctance or refusal to eat unfamiliar foods. This behaviour typically emerges during toddlerhood and peaks between the ages of two and six, which is a critical period for developing food preferences. (Nicklaus & Monnery-Patris, 2018; Zou et al., 2019). While considered a normal part of growth, persistent food neophobia can restrict dietary variety and lead to poor eating habits. Children with high levels of food neophobia often consume fewer fruits, vegetables, and protein-rich foods, preferring sweet, salty, and calorie-dense snacks. (de Almeida et al., 2020). However, such patterns may increase the risk of both undernutrition and overnutrition, negatively affecting growth and development. (Łoboś & Januszewicz, 2019).

Parental influence plays a significant role in the development of children's eating behaviours. Parents and caregivers are typically responsible for the types of foods offered and the overall mealtime environment. Research suggests that children are more likely to accept a wide variety of foods when exposed to them early and repeatedly in a positive setting. Additionally, practices such as using pressure to eat, lack of modelling, or inconsistent feeding routines may further discourage children from trying unfamiliar foods. Kaar et al. (2016) also found that parental practices more strongly influenced children's food neophobia than demographic factors, and that a failure to introduce vegetables early was associated with increased food refusal and weight concerns.

Socioeconomic and environmental factors also contribute to food neophobia. Flight et al. (2003) reported that children from urban areas and higher socioeconomic

* *Corresponding author.*
E-mail address: hazirahjaafar@iium.edu.my

backgrounds exhibited lower levels of food neophobia, likely due to greater exposure to a diverse range of foods. These children were more familiar with various cuisines and more open to trying new foods. In contrast, children with limited exposure, often from rural or lower-income households, were more hesitant to accept unfamiliar foods. While correlations between food neophobia and factors such as cultural diversity, familiarity, and socioeconomic status were low to moderate, they still point to the importance of food exposure and environment in shaping children's dietary behaviours.

In Malaysia, food neophobia among preschool-aged children remains unexplored. While some studies have focused on school-aged populations, preschoolers at a critical stage of developing food preferences have received limited attention. For example, Tan et al. (2024) reported varying levels of food neophobia among children aged 9 to 12 years, yet comparable data on younger children are scarce. Most available evidence is derived from Western populations, which may not reflect Malaysia's multicultural food environment. Hence, this study aimed to assess food neophobia, food consumption, and their associations with sociodemographic factors among preschool children in Kuantan, Pahang.

MATERIALS AND METHODS

Participants

A cross-sectional study was conducted among caregivers of children aged 4 to 6 years, using universal sampling. Data were collected from February to April 2025 at IIUM Educare, Kuantan branch. A total of 64 caregiver-child pairs were invited to participate. Inclusion criteria were children aged 4 to 6 years, while those with chronic illnesses or allergies were excluded.

Sociodemographic Questionnaire

Sociodemographic information of the caregivers included age, gender, race, marital status, educational level, occupation, and household income. Additionally, data related to the preschool children were also collected, including gender, age, age at the introduction of complementary feeding, and duration of breastfeeding.

Anthropometric Measurements

Anthropometric measurements include height, weight, and body mass index (BMI). Height was measured twice using a stadiometer and recorded in centimetres (cm), while weight was measured twice using a digital weighing scale and recorded in kilograms (kg). BMI was calculated using the formula: weight (kg) divided by height in meters squared (m^2). Then, nutritional status was classified according to the World Health Organization (WHO) Child Growth Standards (2008), which categorize them into overweight, obese, stunted, wasted, and severely wasted, depending on the z-score range of growth indicators.

Child Food Neophobia Scale (CFNS) Questionnaire

The Child Food Neophobia Scale (CFNS) questionnaire used in this study was adapted from Luo et al. (2023), (Cronbach's alpha=0.759), which consists of 10 items, probing issues on children's reluctance on unfamiliar food. Each item was rated by caregivers using a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree), based on their child's typical response to unfamiliar foods. The total score ranges from 10 to 70, with higher scores indicating a higher degree of food neophobia. Then, scores were categorised using percentiles representing the degree of food neophobia: < 25th percentile: low food neophobia, 25th to 75th percentile: moderate, and > 75th percentile: high.

Food Frequency Questionnaire

The Food Frequency Questionnaire (FFQ) was adapted from Fatihah et al. (2015) and Noor Hafizah et al. (2019) to assess the frequency of children's consumption of 13 food groups. Frequency options included: 1–3 times per month, once per week, 2–4 times per week, 5–6 times per week, once daily, 2–3 times per day, more than 4 times per day, or never. To improve accuracy, participants were provided with images of standard serving sizes. Caregivers completed the questionnaire based on their child's eating frequency and number of servings consumed. Scoring was based on consumption frequency with reference to portion sizes, and the total score reflected the amount of food consumed across food groups.

Statistical Analysis

Descriptive statistics were used to determine the prevalence of food neophobia, food consumption patterns, nutritional status, and key sociodemographic characteristics (parental education and household income). Independent t-tests and Mann–Whitney U tests were applied to compare nutritional status and food consumption by gender, following normality testing of continuous variables. Associations between food neophobia and parental education, household income, and nutritional status were examined using chi-square tests. All analyses were conducted with SPSS version 20.0,

with statistical significance set at $p < 0.05$.

Ethical Approval

This study received ethical approval from the International Islamic University Malaysia Research Ethics Committee (IREC) (IREC 2025-KAHS/DNS4).

RESULTS

Characteristics of Participants

A total of 64 pairs of caregivers and their preschool children were recruited, with only 44 returned questionnaires, yielding a response rate of 68.8%.

Most caregivers were female (79.5%, n = 35), aged 30–39 years (61.4%, n = 27), had attained higher education (degree/master's/PhD) (72.7%, n = 32), and were employed (93.2%, n = 41). The mean monthly household income was RM 10,519.51 (SD ± 6,768.87), with 29.5% (n = 13) reporting an income of RM 10,000 and above (Table 1).

An equal proportion of boys and girls was recruited. The majority were aged 4 years (52.3%, n = 23), followed by 5 years (27.3%, n = 12) and 6 years (20.5%, n = 9), with a mean age of 4.68 years (SD ± 0.80). Regarding infant feeding practices, most were breastfed for more than 6 months (61.4%, n = 27) and introduced to complementary feeding after 6 months of age (79.5%, n = 35).

Table 1: Sociodemographic characteristics and anthropometric measurements of children (N=44)

Characteristics	**n (%)**	**Mean $\pm$ SD**
Caregivers		
Age		37.25 ± 4.73
21-29 years	2 (4.5%)	
30-39 years	27 (61.4%)	
40-49 years	14 (31.8%)	
50-59 years	1 (2.3%)	
Gender		
Male	9 (20.5%	
Female	35 (79.5%)	
Race		
Malay	44 (100%	
Educational level		
Degree/ Master/ PHD	32 (72.7%)	
Diploma	10 (22.7%)	
SPM	2 (4.5%)	
Occupation Status		
Working	41 (93.2%)	
Self working	1 (2.3%)	
Not working	2 (4.5%)	
Household Income		10519.51 $\pm$ 6768.87
RM 2501 – RM 5000	11 (25.0%)	
RM 5001 – RM 7500	7 (15.9%)	
RM 7501 – RM 10000	10 (22.7%)	
≥ RM 10000	13(29.5%)	
*Missing	3 (6.8%	
Children		
Characteristics	**n (%)**	**Mean $\pm$ SD**
Gender		
Male	21 (47.7%)	
Female	23 (52.3%)	
Age of Children		
4 years old	23 (52.3%)	4.68 ± 0.80
5 years old	12 (27.3%)	
6 years old	9 (20.5%)	
Breastfeeding Period		
Never	2 (4.5%)	3.28 ± 1.10
< 6 months	15 (34.1%)	
> 6 months	27 (61.4%)	
Child Weaning Age		
4-6 months	9 (20.5%)	1.80 ± 0.41
>6 months	35 (79.5%)	
Child's weight (kg)		14.61 ± 3.72
Child's height (cm)		96.12 ± 22.50
BMI ($\mathrm{kg}/\ m^2$)		14.01 ± 4.04
Nutritional Status		
Weight-for-age (WAZ – score)		-0.28 ± 0.59
Normal	40 (90.9%)	

Underweight	3 (6.8%)	
Height-for-age (HAZ – score)		-0.91 ± 0.97
Normal	33 (75.0%)	
Stunting	10 (22.7%)	
BMI-for-age (BAZ – score)		-0.16 ± 0.87
Normal	40 (90.9%)	
Wasting	3 (6.8%)	

Nutritional Status of Preschool Children

Table 1 represents the anthropometric measurements and nutritional status of 44 children. Their average height was 96.12cm (SD±22.50), while the average weight was 14.61kg (SD±3.72). The mean z-scores for weight-for-age, height-for-age, and BMI-for-age were -0.28 (SD±0.50), -0.91 (SD ±0.97) and -0.16 (SD± 0.87), respectively. According to the WHO growth chart, 6.8% (n = 3) of the children were underweight, 22.7% were stunted (n = 10), and 6.8 % (n = 3) were wasted.

Prevalence of Food Neophobia Among Preschool Children

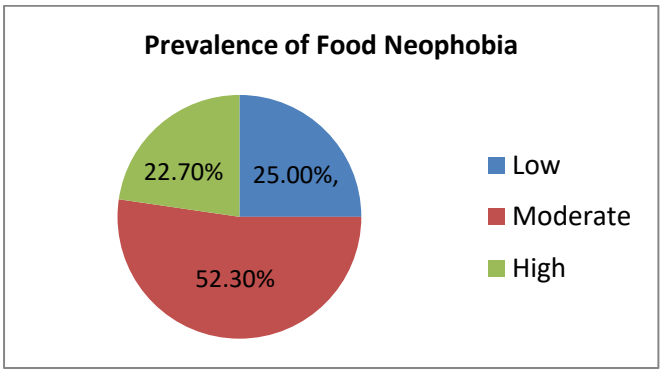


Figure 1: Prevalence of Food Neophobia

The average score for children's food neophobia is 41.09 (SD ±6.24). In this study, 25.0% displayed a low level of food neophobia (n = 11), 52.3% demonstrated a moderate level (n = 23), and 22.7% exhibited a high level of food neophobia (n = 10) (Figure 1).

Food Consumption of Food Neophobic Children

The normality tests conducted prior to statistical analysis indicate that all types of food are not normally distributed. Therefore, the median and interquartile range were reported. Based on Table 2, intake of beverages was the highest among 13 food groups with median 845.50 gram (IQR ± 470.55). Next, cereal and cereals product were the second highest intake with median 844.33 gram (IQR±544.08). Lastly, nuts and bread spreads reported for lowest intake with median 0 (IQR±13.30) and 5.32 gram (IQR±13.15) respectively.

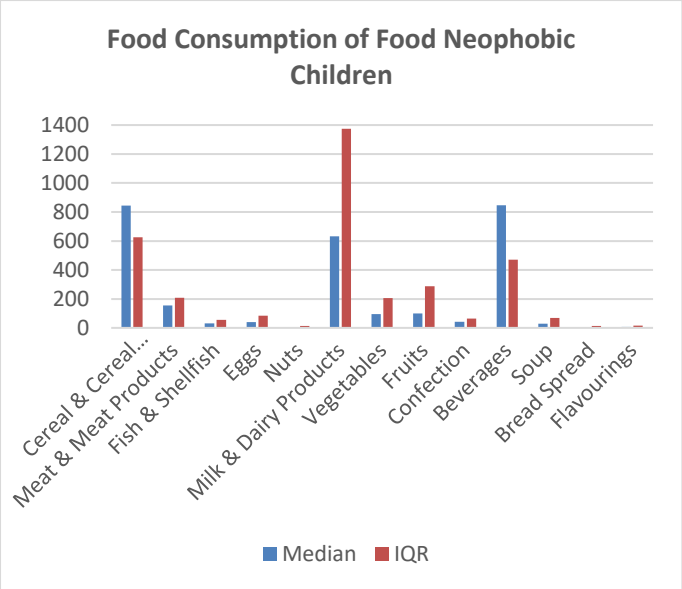


Figure 2: Food Consumption of Food Neophobic Children (N=44)

Factors Influencing Food Neophobia

Tables 3 and 4 present the results using chi-square tests of independence to examine the association between sociodemographic and nutritional factors with food neophobia in preschool children. Table 2 provides the overall associations, while Table 4 offers a detailed distribution of food neophobia levels (low, moderate, high) across categories such as gender, age, parental education, household income, parental occupation, breastfeeding duration, age of complementary feeding, and nutritional status. Although some variations in distribution were observed, for instance, a higher proportion of girls were classified with high food neophobia compared to boys, none of these associations reached statistical significance (all p-values> 0.05). These findings consistently indicate that food neophobia was not significantly associated with sociodemographic or nutritional factors in this sample.

Table 3: Factors Influencing Food Neophobia (N=44)

Factor	P - value
Gender	0.150
Age	1.000
Parental Education	0.188
Household Income	0.468
Occupation Status	0.535
Period of breastfeeding	0.292
Introduction to weaning	0.629
Nutritional Status	0.631

Association Between Food Neophobia and Sociodemographic Status

No significant association was found between food neophobia and any of the sociodemographic characteristics assessed (Table 4). These included gender (X^2(2, n = 44) = 4.010, p = 0.150), age (X^2 (4, n = 44) = 0.107, p = 1.000), parental education level (X^2 (6, n = 44) = 8.825, p = 0.188), household income (X^2(4, n = 44) = 3.842, p = 0.468), parental employment status (X^2(4, n = 44) = 4.174, p = 0.535), duration of breastfeeding (X^2(8, n = 44) = 9.611, p = 0.292), and weaning age (X^2(2, n = 44) = 0.998, p = 0.629). Across all categories, the prevalence of food neophobia did not differ significantly, indicating that neither sociodemographic characteristics nor early feeding practices were associated with the level of food neophobia among the preschool children in this sample.

Association Between Food Neophobia and Nutritional Status

No significant association was found between food neophobia and the nutritional status of preschool children (Table 4). This included weight-for-age z-score (WAZ) (X^2(4, n = 43) = 1.472, p = 0.838), height-for-age z-score (HAZ) (X^2(8, n = 43) = 10.326, p = 0.244), and BMI-for-age z-score (BAZ) (X^2(6, n = 43) = 4.680, p = 0.631). Across all nutritional indicators, the prevalence of low, moderate, and high food neophobia did not differ significantly, indicating that nutritional status was not related to food neophobia levels in this sample.

DISCUSSION

This study provides an overview of food neophobia among preschool children aged 4 to 6 years in Kuantan, Pahang, and its relationship with sociodemographic factors, nutritional status, and food consumption. Most children in this study exhibited moderate food neophobia, followed by low and high levels. Stunting was the most prevalent nutritional issue, while cereals and cereal-based products were the most frequently consumed food groups. Despite expectations, no significant associations were found between food neophobia and either sociodemographic characteristics or nutritional status.

Stunting remains a key nutritional concern in this population. The prevalence found in this study mirrors national trends reported in the National Health and Morbidity Survey (NHMS) 2016, which found stunting rates of 22.5% among Malay children under five (Logarajan et al., 2023). Previous studies have linked stunting to male gender, rural residence, poor sanitation, and limited access to play materials, while protective factors include maternal employment in the private sector and higher protein intake (Mohd Nasir et al., 2012). Evidence suggests that stunting affects both physical growth and cognitive outcomes, underscoring the importance of dietary diversity in early childhood. Regionally, Pahang's stunting prevalence is among the highest in Peninsular Malaysia (Muhammad Zulfahmi Haron et al., 2023), reflecting the need for targeted interventions addressing both undernutrition and micronutrient deficiencies, particularly vitamin D (Poh Bee Koon et al., 2023).

Table 4: Association Between Food Neophobia and Sociodemographic Status and Nutritional Status

Sociodemographic	Low Food Neophobia	Moderate Food Neophobia	High Food Neophobia	n	X^2-statistic (df)	p-value
Gender						
Male	7 (35.0%)	11(55.0%)	2 (10.0%)	20	4.010	0.150
Female	4 (16.7%)	12(50.0%)	8 (33.3%)	24	(2)	*0.306
Age						
4 years old	6 (26.1%)	12(52.2%)	5 (21.7%)	23	0.107	1.000
5 years old	3 (25.0%)	6 (50.0%)	3 (25.0%)	12	(4)	
6 years old	2 (22.2%)	5 (55.6%)	2 (22.2%)	9		
Parental Education						
SPM	1 (50.0%)	1 (50.0%)	0 (0.0%)	2	8.825	0.188
Diploma	2 (20.0%)	5 (50.0%)	3 (30.0%)	10	(6)	
Degree/ Master	8 (28.6%)	16(57.1%)	4 (14.3%)	28		

PhD	0 (0.0%)	1 (25.0%)	3 (75.0%)	4		
Household Income						
B40	3 (27.3%)	6 (54.5%)	2 (18.2%)	11	3.842	0.468
M40	4 (21.1%)	13(68.4%)	2 (10.5%)	19	(4)	
T20	3 (27.3%)	4 (36.4%)	4 (36.4%)	11		
Working Status						
Working	9 (22.0%)	22(53.7%)	10(24.4%)	41	4.174	0.535
Self-working	1(100.0%)	0 (0.0%)	0 (0.0%)	1	(4)	
Not Working	1 (50.0%)	1 (50.0%)	0 (0.0%)	2		
Breastfeeding Period						
Never	0 (0.0%)	2(100.0%)	0 (0.0%)	2	9.611	0.292
< 1 month	1 (33.3%)	0 (0.0%)	2 (66.7%)	3	(8)	
1 – 3 month	0 (0.0%)	4 (66.7%)	2 (33.3%)	6		
4 – 6 month	1 (16.7%)	4 (66.7%)	1 (16.7%)	6		
> 6 month	9 (33.3%)	13(48.15)	5 (18.5%)	27		
Child Weaning Age						
4 – 6 month	2 (22.2%)	6 (66.7%)	1 (11.1%)	9	0.998	0.629
> 6 month	9 (25.7%)	17(48.6%)	9 (25.7%)	35	(2)	*0.162
Nutritional Status						
WAZ Score						
Normal	10(25.0%)	21(52.5%)	9 (22.5%)	40	1.472	0.838
Underweight	1 (33.3%)	1 (33.3%)	1 (33.3%)	3	(4)	
HAZ Score						
Normal	10(30.3%)	14(42.4%)	9 (27.3%)	33	10.326	0.244
Stunting	1(10.0%)	8 (80.0%)	1(10.0%)	10	(8)	
BAZ Score						
Normal	10(25.0%)	22(55.0%)	8 (20.0%)	40	4.680	0.631
Overweight	0 (0.0%)	0 (0.0%)	0 (0.0%)	0	(6)	
Wasting	1 (33.3%)	0 (0.0%)	2 (66.7%)	3		

* Chi Square Test of Independence

*Value of Cramer's V

This finding of no significant relationship between food neophobia and nutritional status aligns with some prior research, which suggests that food neophobia may be shaped more by behavioural and psychological factors than by demographic or nutritional variables (Johnson et al., 2015). Factors such as early exposure to diverse foods, parental feeding practices, and children's sensory sensitivity have been shown to influence neophobia more directly than income or education levels. While socioeconomic status did not appear to influence neophobia in this sample, previous work indicates it can still shape overall dietary patterns and food availability.

The dietary patterns in this study indicate that highly food-neophobic children consumed fewer vegetables, fruits, and legumes, while favouring familiar items such as beverages and cereals. This trend is consistent with previous studies, which found that higher food neophobia scores are associated with narrower dietary variety and reduced preference for nutrient-dense foods (Russell & Worsley, 2008; Perry et al., 2015; Cooke et al., 2006). Such selective eating behaviours may reduce dietary quality, even if total energy intake remains unaffected, potentially compromising long-term health.

Gender differences in food consumption were notable, with females reporting higher intakes of vegetables, fruits, soups, and confectionery compared to males. This finding is consistent with research showing that girls often demonstrate greater willingness to consume nutrient-rich foods, possibly due to parental feeding practices or cultural norms (Wardle et al., 2005; Barbosa et al., 2023; Predieri et al., 2020). These differences suggest that interventions to improve dietary diversity may need to be gender-sensitive, encouraging acceptance of healthy foods, particularly among boys.

The strengths of this study include its focus on preschool-aged children, a developmental stage when food neophobia is particularly relevant, and its categorisation of neophobia levels, which improves the applicability of findings for parents and educators. However, the relatively small sample size and cross-sectional design limit the generalisability and causal interpretation of results.

Reliance on caregiver self-reports may have introduced recall or social desirability bias, and incomplete sociodemographic data further constrained analysis. Additionally, the study did not examine cultural practices, parental feeding styles, or exposure frequency, which could provide further insight into the development of food neophobia.

The findings, together with existing literature, point to a need for multifaceted interventions that promote dietary diversity, enhance parental engagement in feeding, and address environmental and socio-cultural factors. Future research should incorporate larger and more diverse samples, longitudinal designs, and a broader range of influencing factors to better understand and address food neophobia in Malaysian preschoolers.

CONCLUSION

The findings reveal that most preschool children in Kuantan exhibit moderate levels of food neophobia, with no significant association observed between food neophobia and sociodemographic or nutritional status. However, the presence of stunting and wasting among some children highlights ongoing nutritional concerns that require attention. Caregiver factors such as education level and breastfeeding practices showed limited influence on neophobia, suggesting that behavioural and environmental factors may play a greater role. Nutritionists, dietitians, and early childhood educators should prioritize strategies that promote food acceptance, including repeated exposure to unfamiliar foods, caregiver modelling, and positive reinforcement, to encourage a more balanced diet and reduce neophobic tendencies from an early age.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Barbosa, C., Lopes, C., Costa, A., Warkentin, S., & Oliveira, A. (2023). Parental child-feeding practices at 4 years of age are associated with dietary patterns of 7-year-olds. *Journal of Human Nutrition and Dietetics: The Official Journal of the British Dietetic Association, 36*(4), 1339–1348. https://doi.org/10.1111/jhn.13213

Cooke, L., Carnell, S., & Wardle, J. (2006). Food neophobia and mealtime food consumption in 4–5 year old children. *International Journal of Behavioral Nutrition and Physical Activity, 3*(1), 14. https://doi.org/10.1186/1479-5868-3-14

de Almeida, P. C., Rosane, B. P., Nakano, E. Y., Vasconcelos, I. A. L., Zandonadi, R. P., & Botelho, R. B. A. (2020). Instrument to identify food neophobia in Brazilian children by their caregivers. *Nutrients, 12*(7), 1943. https://doi.org/10.3390/nu12071943

Fatihah, F., Ng, B. K., Hazwanie, H., Norimah, A. K., Shanita, S. N., Ruzita, A. T., & Poh, B. K. (2015). Development and validation of a food frequency questionnaire for dietary intake assessment among multi-ethnic primary school-aged children. *Singapore Medical Journal, 56*(12), 687–694. https://doi.org/10.11622/smedj.2015186

Flight, I., Leppard, P., & Cox, D. N. (2003). Food neophobia and associations with cultural diversity and socio-economic status amongst rural and urban Australian adolescents. *Appetite, 41*(1), 51–59. https://doi.org/10.1016/S0195-6663(03)00039-4

Johnson, S. L., Davies, P. L., Boles, R. E., Gavin, W. J., & Bellows, L. L. (2015). Young children's food neophobia characteristics and sensory behaviors are related to their food intake. *The Journal of Nutrition, 145*(11), 2610–2616. https://doi.org/10.3945/jn.115.214284

Kaar, J. L., Shapiro, A. L. B., Fell, D. M., & Johnson, S. L. (2016). Parental feeding practices, food neophobia, and child food preferences: What combination of factors results in children eating a variety of foods? *Food Quality and Preference, 50,* 57–64. https://doi.org/10.1016/j.foodqual.2016.01.006

Łoboś, P., & Januszewicz, A. (2019). Food neophobia in children. *Pediatric Endocrinology Diabetes and Metabolism, 25*(3), 150–154. https://doi.org/10.5114/pedm.2019.90242

Logarajan, R. D., Mohamed Nor, N., Ibrahim, S., & Said, R. (2023). Social determinants of stunting among Malay children below five years in Malaysia. *Nutrition,* 112030. https://doi.org/10.1016/j.nut.2023.112030

Luo, W., Cai, Q., Chen, R., Yin, Y., Sun, X.-X., Cai, Y.-P., Song, H.-Z., Zhang, Y.-R., & Liao, Y. (2023). The Chinese version of the child food neophobia scale and its reliability and validity in preschool children. *Appetite, 69,* 1–5. https://doi.org/10.1016/j.appet.2023.106468

Mohd Nasir, M. T., Norimah, A. K., Hazizi, A. S., Nurliyana, A. R., Loh, S. H., & Suraya, I. (2012). Child feeding practices, food habits, anthropometric indicators and cognitive performance among preschoolers in Peninsular Malaysia. *Appetite, 58*(2), 525–530. https://doi.org/10.1016/j.appet.2011.12.004

Muhammad Zulfahmi Haron, Rohana Abdul Jalil, Noor Aman Hamid, Mohd Azahadi Omar, & Noor Hashimah Abdullah. (2023). Stunting and its associated factors among children below 5 years old on the East Coast of Peninsular Malaysia: Evidence from the National Health and Morbidity Survey. *The Malaysian Journal of Medical Science, 30*(5), 155–168. https://doi.org/10.21315/mjms2023.30.5.12

Nicklaus, S., & Monnery-Patris, S. (2018). Food neophobia in children and its relationships with parental feeding practices/style. In *Handbook of Eating and Drinking* (pp. 255–286). Springer. https://doi.org/10.1007/978-3-319-75388-1_17

Noor Hafizah, Y., Ang, L. C., Yap, F., Nurul Najwa, W., Cheah, W. L., Ruzita, A. T., Jumuddin, F. A., Koh, D., Lee, J. A. C., Essau, C. A., & Reeves, S. (2019). Validity and reliability of a food frequency questionnaire (FFQ) to assess dietary intake of preschool children. *International Journal of Environmental Research and Public Health, 16*(23), 4722. https://doi.org/10.3390/ijerph16234722

Perry, R. A., Mallan, K. M., Koo, J., Mauch, C. E., Daniels, L. A., & Magarey, A. M. (2015). Food neophobia and its association with diet quality and weight in children aged 24 months: A cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 13. https://doi.org/10.1186/s12966-015-0184-6

Poh, B. K., Wong, J. E., Lee, S. T., Yeo, G. S., Sharif, R., Shanita, S. N., Jamil, N. A., Mei, C., Sameeha, M. J., Koh, D., Lim, N. S. M., Norimah, A. K., Ruzita, A. T., Budin, S. B., Wee, L. H., Tang, S. F., & Khouw, I. (2023). Triple burden of malnutrition among Malaysian children aged 6 months to 12 years: Current findings from SEANUTS II Malaysia. *Public Health Nutrition,* 1–36. https://doi.org/10.1017/S1368980023001153

Predieri, S., Sinesio, F., Monteleone, E., Spinelli, S., Cianciabella, M., Daniele, G. M., Dinnella, C., Gasperi, F., Endrizzi, I., Torri, L., Toschi, T. G., Bendini, A., Pagliarini, E., Cattaneo, C., Di Monaco, R., Vitaglione, P., Condelli, N., & Laureati, M. (2020). Gender, age, geographical area, food neophobia and their relationships with the adherence to the Mediterranean diet: New insights from a large population cross-sectional study. *Nutrients, 12*(6), 1778. https://doi.org/10.3390/nu12061778

Russell, C. G., & Worsley, A. (2008). A population-based study of preschoolers' food neophobia and its associations with food preferences. *Journal of Nutrition Education and Behavior, 40*(1), 11–19. https://doi.org/10.1016/j.jneb.2007.03.007

Tan, X., Shukor, A., & Soh, K. G. (2024). Visual cues, liking, and emotional responses: What combination of factors result in the willingness to eat vegetables among children with food neophobia? *Foods, 13*(20), 3294. https://doi.org/10.3390/foods13203294

Wardle, J., Carnell, S., & Cooke, L. (2005). Parental control over feeding and children's fruit and vegetable intake: How are they related? *Journal of the American Dietetic Association, 105*(2), 227–232. https://doi.org/10.1016/j.jada.2004.12.009

Zou, J., Liu, Y., Yang, Q., Liu, H., Luo, J., Ouyang, Y., Wang, J., & Lin, Q. (2019). Cross-cultural adaption and validation of the Chinese version of the child food neophobia scale. *BMJ Open, 9*(8), e026729. https://doi.org/10.1136/bmjopen-2018-026729

ORIGINAL ARTICLE

Evaluating the Association Between Dairy Consumption, Acne Development and Quality of Life Among University Students

Adrina Munirah Shariman[1], Nuraniza Azahari[1,2*], Noraishah Mohamed Nor[1,2], Aflah Afandi[1,2]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25000 Kuantan, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
Acne vulgaris; dairy consumption; university students; quality of life

ABSTRACT

Background: Acne vulgaris is one of the most prevalent dermatological conditions among adolescents and young adults, often impairing self-esteem, social relationships, and quality of life. Dietary factors, particularly dairy intake, have been suggested to contribute to acne, although findings remain inconsistent. This study aimed to evaluate the association between dairy consumption, acne severity, and quality of life among the students of the International Islamic University Malaysia (IIUM). **Methods:** A cross-sectional study was conducted among 377 undergraduate students from three IIUM campuses. Data were collected using a self-administered questionnaire comprising a modified Food Frequency Questionnaire (FFQ) for dairy intake, the Global Acne Grading System (GAGS) for acne severity, and the Acne-Specific Quality of Life (Acne-QoL) questionnaire. Spearman's rank correlation was performed to determine associations between variables. **Results:** The mean daily dairy intake among students was 470.68 ± 183.88 g/day. Yoghurt and fresh milk were the most consumed dairy products. No significant association was found between overall dairy intake, fresh milk, yoghurt, or cheese with acne severity. However, commercial milk, powdered milk, and evaporated milk showed a significant positive correlation with acne severity ($p < 0.005$). Regarding quality of life, acne severity was negatively correlated with all four Acne-QoL domains (self-perception, role-social, role-emotional, and acne symptoms), with the strongest correlation observed in the acne symptoms domain. **Conclusion:** This study suggests that certain forms of dairy, particularly commercial, powdered, and evaporated milk, may contribute to acne severity among university students. Acne was also found to adversely affect multiple domains of quality of life. These findings highlight the importance of dietary counselling and holistic acne management strategies for young adults. Further research considering additional lifestyle and genetic factors is warranted.

INTRODUCTION

Acne vulgaris is a prevalent chronic inflammatory skin disease that commonly affects adolescents and young adults worldwide. It is characterised by comedones, papules, pustules, and nodules, which develop due to follicular hyperkeratinisation, increased sebum production, colonisation by Cutibacterium acnes, and inflammation (Xu et al., 2024). While acne is often considered a temporary cosmetic issue, it can have significant consequences on self-esteem, social interaction, and psychological well-being, particularly among university students navigating a critical stage of academic and personal development (Tan et al., 2022; Natali et al., 2023). Several factors contribute to acne pathogenesis, including genetics, hormones, stress, and environmental influences.

Diet has increasingly been recognised as a modifiable risk factor. High glycaemic index diets, refined sugars, and unhealthy fat intake have been associated with increased acne prevalence (Huang et al., 2018; Podgórska et al., 2021). In particular, dairy products have received growing attention due to their hormonal content and bioactive compounds. Milk and its derivatives contain insulin-like growth factor-1 (IGF-1), androgens, and progesterone, which can stimulate sebaceous gland activity, keratinocyte proliferation, and sebum production, thereby contributing to acne development (Aalemi et al., 2019).

The literature, however, remains inconclusive. Meta-analyses suggest a significant association between dairy intake, especially milk and acne (Aghasi et al., 2019), while other studies report no such correlation (Juhl et al., 2018). Differences in the study design, sample characteristics, and dietary patterns may account for these inconsistencies. Furthermore, different dairy types

* *Corresponding author.*
E-mail address: nuraniza@iium.edu.my

appearto exert varying effects. For instance, low-fat milk has been more strongly linked with acne than whole milk, potentially due to higher IGF-1 levels and increased insulin response (González-Mondragónet al., 2022). Conversely, fermented dairy products such as yoghurt may have beneficial effects on the skin due to their probiotic content, although results remain mixed (Aslam & Younas, 2020).

Beyond its physical manifestations, acne significantly impairs quality of life. Affected individuals often report poor self-perception, emotional distress, and social withdrawal, which can lead to reduced academic performance and psychosocial well-being (Mehra et al., 2023; Pourramzani et al., 2024). University students, in particular, are vulnerable due to lifestyle factors such as stress, irregular sleep, and changing eating habits, which may exacerbate acne and its psychological impact. Despite this, limited research has explored the link between acne and quality of life in Malaysian populations, where dietary habits reflect both traditional and Western influences.

Given the scarcity of local evidence and the inconsistent findings globally, further research is needed to clarify the role of dairy consumption in acne development and its subsequent effect on students' well-being. Therefore, this study aimed to evaluate the association between dairy intake and acne severity, as well as to examine how acne influences the quality of life among students at the International Islamic University Malaysia (IIUM).

MATERIALS AND METHODS

Study Design and Setting

A cross-sectional study was conducted among undergraduate students at the International Islamic University Malaysia (IIUM). Data collection was conducted across three campuses: Gombak (Kuala Lumpur), Kuantan (Pahang), and Pagoh (Johor).

Participants and Sampling

The study population comprised undergraduate students from various kulliyyahs and academic years. Postgraduate, foundation students, and those undergoing medical treatment for acne (e.g., antibiotics, retinoids, hormonal therapy), as well as students with lactose intolerance or dairy allergies, were excluded. Using Krejcie and Morgan's (1970) sampling table, a representative sample of 377 students was recruited through convenience sampling.

Ethical Approval

Ethical approval was obtained from the IIUM Research Ethical Committee (IREC) and the Kulliyyah Postgraduate and Research Committee (KPGRC). Written informed consent was obtained from all participants, and confidentiality was maintained throughout the study.

Data Collection Instruments

Data were gathered through a self-administered online questionnaire consisting of four sections:

1. Sociodemographic Information – age, gender, kulliyyah, campus, and year of study.
2. Dairy Consumption – assessed using a modified Food Frequency Questionnaire (FFQ) adapted from the Malaysian Adult Nutrition Survey 2014. The FFQ captured the frequency and portion sizes of six main dairy products: fresh milk, commercial milk, powdered milk, evaporated milk, yoghurt, and cheese. Mean daily intake (grams/day) was calculated using standard conversion factors (Norimah, et al., 2008).
3. Acne Severity – measured using the Global Acne Grading System (GAGS), which scores acne lesions across six anatomical regions (forehead, cheeks, nose, chin, chest, and upper back). Scores range from 0 (none) to >39 (very severe), and severity is categorised as mild, moderate, severe, or very severe (Gorelick et al., 2015). 4. Quality of Life – assessed using the validated Acne-Specific Quality of Life (Acne-QoL) questionnaire, covering four domains: self-perception, role-social, role-emotional, and acne symptoms. Items were scored on a 7-point Likert scale, with higher scores indicating less impairment (Gupta, 1998).

Pilot Study

A pilot test with 30 students was conducted to assess the reliability and clarity of the instruments. Cronbach's alpha values were acceptable for the Acne-QoL ($\alpha = 0.90$), while the FFQ and GAGS showed lower internal consistency ($\alpha = 0.30$). The relatively low value could be explained by the sensitivity of Cronbach's alpha to both sample size and the number of items, since the FFQ and GAGS consist of only a limited number of questions. No modifications were made, as both are widely validated tools in the fields of nutrition and dermatology research.

Statistical Analysis

Data were analysed using SPSS version 29. Descriptive statistics were used for sociodemographic data, dairy intake, acne severity, and quality of life scores. Normality testing indicated skewed distribution; thus, non-parametric tests were applied. Spearman's rank correlation was used to determine the association between dairy intake and acne severity, and the association between acne severity and quality of life domains. A p-value of <0.05 was considered statistically significant

RESULTS

Sociodemographic Characteristics

A total of 377 students participated in the study. The majority were female (82.5%), Malay (98.4%), and aged between 21–23 years (71.1%). Most were from the Kuantan campus (64.7%) and enrolled in the Kulliyyah of Allied Health Sciences (38.2%).

Dairy Consumption

The mean total dairy intake was 470.68 ± 183.88 g/day. Among the six categories, yoghurt (57.46 ± 103.21 g/day) and fresh milk (53.69 ± 96.83 g/day) were the most consumed items, whereas cheese (10.94 ± 20.80 g/day) had the lowest intake (Table 1).

Table 1: Mean Daily Dairy Intake (g/day)

Dairy product	Mean ± SD
Fresh milk	53.69 ± 96.83
Commercial milk	23.14 ± 39.63
Yoghurt	57.46 ± 103.21
Powdered milk	15.51 ± 48.18
Evaporated milk	21.37 ± 43.65
Cheese	10.94 ± 20.80
Total dairy intake	470.68 ± 183.88

Acne Severity

The mean GAGS score was 10.42 ± 6.97, which falls within the mild acne category. The right cheek had the highest mean score (2.30 ± 2.21), followed closely by the left cheek (2.25 ± 2.15). The chin and nose showed relatively lower mean values (1.06 ± 1.15 and 0.90 ± 0.65, respectively) (Table 2).

Table 2: Descriptive analysis of Global Acne Grading System (GAGS)

Region	Mean ± SD
Forehead	1.82 ± 2.07
Right cheek	2.30 ± 2.21
Left cheek	2.25 ± 2.15
Nose	0.90 ± 0.65
Chin	1.06 ± 1.15
Chest and upper back	2.08 ± 2.80
Total acne severity score	10.42 ± 6.97

Quality of Life

The Acne-QoL scores indicated that acne had the most impact on the acne symptoms domain (20.90 ± 5.63), followed by self-perception (19.91 ± 8.50), role-emotional (19.73 ± 8.73), and role-social (17.92 ± 6.73).

Association Between Dairy Intake and Acne Severity

Spearman's correlation showed no significant association between total dairy intake, fresh milk, yoghurt, or cheese with acne severity. However, commercial milk (r = 0.103, p = 0.045), powdered milk (r = 0.114, p = 0.026), and evaporated milk (r = 0.173, p = 0.001) were significantly and positively correlated with acne severity (Table 3).

Table 3: Spearman correlation between dairy intake and acne severity

Dairy product	r	p-value
Fresh milk	0.086	0.095
Commercial milk	0.103	0.045*
Yoghurt	0.033	0.528
Powdered milk	0.114	0.026*
Evaporated milk	0.173	0.001*
Cheese	0.080	0.123
Total dairy intake	0.082	0.110

Association Between Acne Severity and Quality of Life

Acne severity was negatively correlated with all Acne-QoL domains. There was a statistically significant moderate negative correlation between acne severity and the domain self-perception (r = -0.456, p < 0.001), the role emotional domain (r = -0.454, p < 0.001), the role social domain (r = -0.369, p < 0.001), and the acne symptoms domain (r = -0.602, p < 0.001) (Table 4).

Table 4: Spearman correlation between acne severity and Acne-QoL domains

Dairy product	r	p-value
Self-perception	-0.456	0.001*
Role emotional	-0.454	0.001*
Role social	-0.369	0.001*
Acne symptoms	-0.602	0.001*

DISCUSSION

Dairy Consumption and Acne Severity

The mean daily dairy intake among participants was comparable to values reported in previous Malaysian dietary surveys (MOH, 2014), suggesting that students maintain a moderate level of dairy consumption. Consistent with prior studies, yoghurt and fresh milk were the most consumed items (Aslam & Younas, 2020). Despite the high intake, neither fresh milk nor yoghurt demonstrated significant associations with acne severity in this study. This aligns with findings from Juhl et al. (2018), who reported no direct link between overall milk intake and acne among young adults in Denmark. Additionally, according to the Malaysian Dietary Guidelines (MDG 2020), adults are encouraged to consume 2 to 3 servings of milk and milk products daily, equivalent to approximately 400 to 600 g per day of dairy foods such as milk, yoghurt, or cheese (Ministry of Health Malaysia, 2020).

Interestingly, only specific dairy subtypes, which are commercial milk, powdered milk, and evaporated milk, were associated with acne severity. These results echo those of Aghasi et al. (2019), who found that skimmed and processed milk products were more strongly correlated with acne compared to whole milk. One possible explanation lies in the processing methods of these dairy products. Powdered milk and evaporated milk undergo high-temperature processing, which may alter the protein structure and increase insulinotropic effects, thereby enhancing insulin-like growth factor-1 (IGF-1) activity. Elevated IGF-1 stimulates sebaceous gland proliferation and sebum production, both of which are key contributors to acne pathogenesis.

Commercially produced milk, often fortified and homogenised, may also contain residual hormones and higher levels of whey proteins, which have been shown to elevate insulin and IGF-1 responses (Aghasi et al., 2019). By contrast, yoghurt, as a fermented product, may exert a protective effect due to the presence of probiotics, which modulate gut microbiota and systemic inflammation. The lack of significant association between yoghurt and acne severity in this study could therefore be attributed to its counteractive properties, although further research is warranted.

Acne Severity and Quality of Life

Consistent with previous findings (Mehra et al., 2023; Pourramzani et al., 2024), this study found that acne severity was negatively correlated with all domains of the Acne-QoL. The strongest correlation was observed with acne symptoms, indicating that the physical manifestations of acne were perceived as the most distressing aspect for students. This is consistent with Natali et al. (2023), who highlighted that visible lesions and scarring strongly impair quality of life in young adults.

The negative correlation with self-perception and role-social domains underscores the impact of acne on self-image and social relationships. Students with higher acne severity reported poorer confidence, avoidance of social gatherings, and heightened self-consciousness, findings that align with reports from Europe and the Middle East (Machala-Cwikla et al., 2024; Aalemi et al., 2019). Additionally, the significant correlation with the role-emotional domain suggests that acne contributes to mood disturbances, irritability, and feelings of embarrassment, further emphasising the multifaceted burden of the condition.

The results of this study partially support the hypothesis that dairy intake is associated with acne, though only for specific dairy products. While several studies in Western populations have linked all forms of milk to acne risk (Aghasi et al., 2019; González-Mondragón et al., 2022), the present findings suggest that Malaysian students may respond differently, potentially due to differences in dietary habits, dairy processing methods, and genetic predispositions. The higher consumption of yoghurt and traditional dairy products in Malaysia may mitigate acne severity compared to populations with higher reliance on processed dairy.

The quality-of-life findings are consistent across international literature, reinforcing the global nature of acne's psychosocial burden. However, the magnitude of impairment may vary depending on cultural perceptions of beauty and societal pressure. In collectivist cultures such as Malaysia, where physical appearance plays a role in social acceptance, acne may disproportionately affect students' self-esteem and social functioning (Tan et al., 2022).

Strengths and Limitations

A key strength of this study lies in its relatively large sample size (N = 377), covering students from three different campuses, which enhances representativeness. The use of

validated tools (FFQ, GAGS, and Acne-QoL) also strengthens reliability. Importantly, the study distinguishes between different types of dairy, providing nuanced insights often overlooked in previous research.

However, limitations must be acknowledged. The cross-sectional design prevents causal inference, as the temporal relationship between dairy intake and acne cannot be established. Dietary data were self-reported, which may introduce recall bias. Additionally, other dietary and lifestyle factors such as glycaemic load, stress, and sleep patterns were not measured, though they may confound the observed associations. Finally, the predominance of female participants (82.5%) may limit generalisability to male students.

CONCLUSION

This study found that while overall dairy consumption was not significantly associated with acne severity among IIUM students, specific processed dairy products namely commercial milk, powdered milk, and evaporated milk showed a positive correlation with acne severity. In contrast, fresh milk, yoghurt, and cheese demonstrated no significant associations. Additionally, acne severity was found to negatively impact students' quality of life across all domains, with the strongest effect observed in the acne symptoms domain, followed by self-perception, social roles, and emotional well-being.

These findings suggest that dietary counselling for acne management should emphasise moderation of processed dairy products, while recognising that fermented products such as yoghurt may not exert the same adverse effects. Furthermore, acne management should adopt a holistic approach, addressing both physical symptoms and the psychological burden experienced by students.

ACKNOWLEDGEMENT

The authors would like to thank the Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM), for their support in facilitating this study. Special appreciation is extended to all participating students for their cooperation and valuable contributions.

REFERENCES

Aalemi, A. K., Anwar, I., & Chen, H. (2019). Dairy consumption and acne: A case control study in kabul, afghanistan. Clinical, Cosmetic and Investigational Dermatology, 12, 481–487. https://doi.org/10.2147/CCID.S195191

Aghasi, M., Golzarand, M., Shab-Bidar, S., Aminianfar, A., Omidian, M., & Taheri, F. (2019). Dairy intake and acne development: A meta-analysis of observational studies. *Clinical Nutrition,* 38(3), 1067-1075 https://doi.org/10.1016/j.clnu.2018.04.015

Aslam, M., & Younas, M. (2020). Association Between Dietary Practices and Acne in Young Adults; A Review. Journal of Dermatology Research Reviews & Reports, 1–8. https://doi.org/10.47363/JDMRS/2020(1)114

González-Mondragón, E. A., Ganoza-Granados, L. D. C., Toledo-Bahena, M. E., Valencia-Herrera, A. M., Duarte-Abdala, M. R., Camargo-Sánchez, K. A., & Mena-Cedillos, C. A. (2022). Acne and diet: a review of pathogenic mechanisms. Boletin Medico Del Hospital Infantil de Mexico, 79(2), 83–90. https://doi.org/10.24875/BMHIM.21000088

Gorelick, J., Daniels, S. R., Kawata, A. K., Degboe, A., Wilcox, T. K., Burk, C. T., & Douse-Dean, T. (2015). Acne-Related Quality of Life Among Female Adults of Different Races/Ethnicities. *Journal of the Dermatology Nurses' Association*, *7*(3), 154–162. https://doi.org/10.1097/JDN.0000000000000129

Gupta, M. A., Johnson, A. M., & Gupta, A. K. (1998). The development of an Acne Quality of Life scale: reliability, validity, and relation to clinical severity. British Journal of Dermatology, 139(3), 435–440.

Huang, X., Zhang, J., Li, J., Zhao, S., Xiao, Y., Huang, Y., Jing, D., Chen, L., Zhang, X., Su, J., Kuang, Y., Zhu, W., Chen, M., Chen, X., & Shen, M. (2018). Daily Intake of Soft Drinks and Moderate-to-Severe Acne Vulgaris in Chinese Adolescents Article In Press The Journal Of Pediatrics • www.jpeds.com. The Journal of Pediatrics. https://doi.org/10.1016/j.jpeds

Juhl, C. R., Bergholdt, H. K., Miller, I. M., Jemec, G. B., Kanters, J. K., Ellervik, C., & Johansen, J. D. (2018). Dairy intake and acne vulgaris: A systematic review and meta-analysis of observational studies. *Nutrients, 10*(8), 1049. https://doi.org/10.3390/nu10081049

Machała-Ćwikła, E., Ćwikła, P., Zdziebło, P., Łapińska, U., Machała, K., & Machała, D. (2024). Acne vulgaris in the age of social media. Our Dermatology Online, 15(e), e42–e42. https://doi.org/10.7241/ourd.2024e.42

Mehra, A., Kamat, D., Narang, T., Avasthi, A., Sahoo, S., & Grover, S. (2023). Psychological Morbidity and Body Dysmorphic Disorder among Patients with Acne. Journal of Mental Health and Human Behaviour, 28(2), 154–161. https://doi.org/10.4103/jmhhb.jmhhb_123_23

Ministry of Health Malaysia. (2020). *Malaysian Dietary Guidelines 2020*. Putrajaya: National Coordinating Committee on Food and Nutrition, Ministry of Health Malaysia.

Ministry of Health Malaysia (MOH). (2014). *Malaysian Adult Nutrition Survey (MANS) 2014: Volume 1:*

Methodology and General Findings. Institute for Public Health, Ministry of Health Malaysia.

Natali, O., Ayu Athifa Serena, P., & Sari Mutia, M. (2023). The Relationship Between the Severity of Acne Vulgaris and the Quality of Life of Prima Indonesia University Medical Faculty Students. International Journal of Social Health, 2(11), 846–855. https://doi.org/10.58860/ijsh.v2i11.123

Norimah AK, Safiah M, Jamal K, Haslinda S, Zuhaida H, Rohida S, et al. Food consumption patterns: findings from the Malaysian Adult Nutrition Survey (MANS). Mal J Nutr. 2008;14:25–39.

Podgórska, A., Puścion-Jakubik, A., Markiewicz-żukowska, R., Gromkowska-Kępka, K. J., & Socha, K. (2021). Acne vulgaris and intake of selected dietary nutrients— a summary of information. In Healthcare (Switzerland) (Vol. 9, Issue 6). MDPI AG. https://doi.org/10.3390/healthcare9060668

Pourramzani, A., Nejad, K. G., & Ghavami, S. S. (2024). Evaluation of the feelings of loneliness and suicidal thoughts in individuals with acne vulgaris. Iranian Journal of Dermatology, 27(2), 67–74. https://doi.org/10.22034/ijd.2024.421679.1768

Tan, J., Beissert, S., Cook-Bolden, F., Chavda, R., Harper, J., Hebert, A., Lain, E., Layton, A., Rocha, M., Weiss, J., & Dréno, B. (2022). Evaluation of Psychological Wellbeing and Social Impact of Combined Facial and Truncal Acne: a Multi national, Mixed-Methods Study. Dermatology and Therapy, 12(8), 1847–1858. https://doi.org/10.1007/s13555-022-00768-0

Xu, W., Zhang, Y., Sinaki, D. G., Zhang, Z., Tang, Y., & Chen, Y. (2024). Acne-induced pathological scars: pathophysiology and current treatments. In Burns and Trauma (Vol. 12). Oxford University Press. https://doi.org/10.1093/burnst/tkad060

REVIEW ARTICLE

Meal Replacement Plans in Managing Adolescent Obesity: A Systematic Literature Review

Ainul Mardhiah Irfan[1], Noraishah Mohamed Nor[1,2*], Roszanadia Rusali[1,2], & Noor Suzana Osman[1,2]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
Adolescent obesity; Meal replacement plans; Weight management; Sustainability; Safety

ABSTRACT

Background: Adolescent obesity remains a significant public health issue, with rising prevalence and long-term health risks. While lifestyle modifications are the primary treatment, their effectiveness is often limited. Meal Replacement Plans (MRPs) have emerged as a potential alternative, though evidence in adolescents remains unclear. **Objective:** This systematic review aimed to assess the effectiveness, sustainability, and safety of MRPs in managing obesity among adolescents. **Method:** A systematic search was conducted across PubMed, Scopus, and Cochrane Library for studies published between 2013 and 2024. Eligible studies included randomised controlled trials, pilot trials, and cohort studies involving adolescents aged 10–18 years who were overweight or obese. Data on intervention design, weight loss outcomes, sustainability, and side effects were extracted and synthesised narratively. **Results:** Seven studies with sample sizes ranging from 30 to 2,825 adolescents were included. MRPs, delivered as very-low-energy diets or partial replacement regimens, consistently produced short-term weight reduction (3.4–6.3% BMI decrease within 4–12 weeks). Adjunctive supports, including caregiver involvement, pharmacologic agents (e.g., exenatide), and financial incentives, enhanced outcomes. However, long-term sustainability was mixed; several trials showed weight regain after 12 months, while caregiver engagement and digital health tools demonstrated potential for maintaining weight loss. Side effects were generally mild (e.g., hunger, fatigue, nausea) and transient, with no adverse impact on growth reported. **Conclusion:** MRPs are effective for short-term weight reduction in adolescents with obesity and may be most beneficial when integrated with behavioral therapy, caregiver participation, or pharmacologic support. Their long-term sustainability remains uncertain, underscoring the need for extended follow-up studies.

INTRODUCTION

Background

Adolescence is defined as the period between 10 and 19 years of age, and is a crucial developmental phase characterised by rapid physical growth, psychological development, and increasing independence in making health-related decisions (World Health Organization, 2024). At this age, unhealthy eating patterns such as meal skipping, with increased sedentary behaviour and reduced physical activity often begin, contributing to excessive weight gain. Together with hormonal and psychosocial changes, these lifestyle patterns heighten the risk of overweight and obesity (De Lorenzo et al., 2019; Yamada et al., 2023).

Although not classified as a disease in itself, obesity is a major risk factor for chronic conditions. Among children and adolescents aged 5–19 years, the World Health Organization (WHO) defines overweight and obesity using the BMI-for-age chart: overweight is above +1 SD and obesity above +2 SD relative to the WHO Growth Reference median (World Health Organization, 2024). For adults (≥18 years), cut-offs are ≥25 kg/m² for overweight and ≥30 kg/m² for obesity. At the same time, ASEAN-specific thresholds are slightly lower (≥23 kg/m² for overweight and ≥27.5 kg/m² for obesity), which apply to adults (CPG for the Management of Obesity, 2023).

The global prevalence of adolescent obesity has risen dramatically. According to the WHO (2024), obesity among adolescents has quadrupled since 1990. In 2022,

[1*] *Corresponding author.*
E-mail address: ishah@iium.edu.com

Over 390 million children and adolescents (5–19 years) were overweight, of whom 160 million were obese. In Malaysia, the National Health and Morbidity Survey (NHMS) 2022 reported overweight and obesity rates of 16.2% and 14.3%, respectively, among adolescents (Institute for Public Health (IPH), 2022). These figures signal a pressing public health concern requiring urgent action.

Obesity results from a persistent energy imbalance, where energy intake exceeds energy expenditure, leading to the accumulation of fat (Oussaada et al., 2019). For adolescents, this is often linked to high-calorie diets, sedentary lifestyles, and environmental or genetic predispositions. If untreated, obesity can lead to serious health complications, including type 2 diabetes mellitus (T2DM), hypertension, dyslipidemia, cardiovascular disease, stroke, sleep apnoea, and certain cancers (Evans et al., 2023; Lee, 2015).

Lifestyle interventions, such as dietary changes, physical activity, behavioural therapy, and family- or school-based programs, are the first-line treatment for adolescent obesity. While widely implemented, these interventions often have limited effectiveness in severely obese adolescents (Astbury et al., 2019). One promising dietary approach is the use of Meal Replacement Plans (MRPs), which involve substituting one or more daily meals with nutritionally balanced products such as shakes, bars, or prepared meals. MRPs simplify calorie control and are typically rich in protein and fibre to promote satiety (Astbury et al., 2019).

While commonly used in adults, emerging evidence suggests potential effectiveness in adolescents, particularly those with severe obesity or poor response to conventional lifestyle strategies. However, concerns remain regarding long-term sustainability, nutrient adequacy, and safety in this age group. Thus, this systematic review aims to evaluate the efficacy, safety, and sustainability of MRPs in the management of adolescent obesity.

METHODOLOGY

A systematic literature review (SLR) was conducted to synthesise evidence on meal replacement plans (MRPs) for adolescent obesity, guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). This design was chosen for its methodological rigour, ability to appraise study quality, and stronger evidence synthesis compared to scoping reviews (Mohamed Shaffril et al., 2021; Alexander, 2020).

A comprehensive search was conducted in PubMed, Scopus, and the Cochrane Library, utilising Boolean operators (AND, OR, NOT) and Medical Subject Headings (MeSH) to refine the retrieval. Search terms were structured using the Population, Intervention, Comparison, Outcomes, and Study Design (PICOS) framework (Table 1) (Amir-Behghadami & Janati, 2020). Table 2 describes the inclusion and exclusion criteria for the review. Eligible studies included adolescents aged 10–18 years with obesity (BMI ≥35 kg/m²) and interventions involving MRPs. Outcomes of interest were weight reduction, sustainability, and side effects. Cross-sectional, case-control, cohort studies, randomised controlled trials (RCTs), and quasi-experimental studies were included, while reviews, opinion papers, animal studies, and studies with unrelated outcomes were excluded. Articles were screened for titles, abstracts, and full texts, independently by the researcher, with duplicates removed using Mendeley. Only studies meeting the inclusion criteria were qualitatively synthesized.

Table 1: Search Strings

Category	Search Term
Population	(“Obes*” OR “Overweight” OR “Excessive weight” OR “Adiposity”) AND (“Adolescen*” OR “Child*” OR “Teen*” OR “Youth”) OR “Adolescent obesity” OR “Youth obesity” OR “Teen obesity”
Intervention	“Meal replacement*” OR {“Meal replacement” AND (“plan*” OR “therap*)} OR “MRP” OR “MRT” OR "Diet replacement" OR "Partial meal replacement" OR "Liquid diet" OR "Formula diet" OR "Nutritional replacement" OR "Dietary replacement therapy"
Comparison	None
Outcomes	("Reduced BMI" OR "BMI decrease" OR "Weight reduction" OR "Weight loss" OR "Body composition" OR "Fat loss" OR "Body weight" OR "Weight control" OR "Healthy weight" OR "Weight management")
Study design	“Cross-sectional” OR “Case-Control” OR “Cohort” OR “Randomised Controlled Trial” OR “RCT” OR “Cross-sectional Study” OR “Quasi-Experimental Studies”

Table 2: Inclusion & Exclusion Criteria

Elements	Inclusion	Exclusion
Population	Adolescents (10-18 years old) Obesity (BMI ≥35 kg/m²) English articles only Articles until 2025	Animal studies
Intervention	Meal replacement plans	Non-meal replacement dietary interventions
Comparison	None	
Outcomes	Weight reduction (total or percentage)	Not related to obesity reduction or weight management
Study design	Cross-sectional Studies Case-Cohort Studies Cohort Studies Randomized Controlled Trials (RCTs) Quasi-Experimental Studies	Opinion papers Review articles

The methodological quality of the included studies was assessed using the Quality Criteria Checklist for Primary Research (QCCPR) (Academy of Nutrition and Dietetics, n.d.). Each study was rated as positive (+), neutral (Ø), or negative (–) based on its validity and relevance.

Data were extracted for study design, participants, interventions, outcomes, follow-up, sustainability, side effects, and limitations. The findings were summarised in a tabular format, adapted from Schünemann et al. (2024).

This review was registered with PROSPERO (ID: CRD42025634588). Ethical approval was not required, as no human or animal participants were involved.

RESULTS

Figure 1 shows the PRISMA flow diagram, which identifies 197 studies and includes seven studies for final review. The seven studies (Fox et al., 2021; Fox et al., 2016; Gow et al., 2024; Gross et al., 2024; Lei et al., 2020; Phan et al., 2013; Xanthopoulos et al., 2013) met the inclusion criteria, with sample sizes ranging from 30 to 2,825 adolescents. Study designs included randomized controlled trials, cohort studies, and pilot interventions, conducted in diverse populations across multiple countries. Interventions primarily consisted of very-low-energy diets (VLED), total meal replacement plans, or partial replacement regimens combined with standard dietary advice. All included studies were appraised as having overall positive methodological quality, supporting their suitability for synthesis.

All studies met no fewer than two out of four relevance criteria and were overall quality assessed as positive (+). Validity scores varied between 6/10 and 10/10, and the highest quality studies were characterised by methodological rigour at the design, measurement, and reporting stages. Commonly unmet criteria included non-blindedness, inadequate reporting of sources of funding, and inadequate reporting on reasons for participant withdrawal. In general, the studies included have a good level of evidence in support of meal replacement therapy for adolescent obesity intervention, at least in the short term. Nevertheless, the variation in long-term efficacy and differences in methodology point to some needed improvements.

Across all studies, MRPs consistently produced short-term reductions in BMI or body weight. Reported decreases ranged between 3.4% and 6.3% over intervention periods of 4–12 weeks (Berkowitz et al., 2011; Khayutin et al., 2023). Adjunctive strategies, such as caregiver involvement, pharmacological support (e.g., exenatide), and financial incentives, further enhanced weight outcomes. However, the long-term sustainability of these reductions was mixed. Several trials reported partial or complete weight regain after 12 months; however, programs incorporating caregiver engagement or digital health monitoring demonstrated better maintenance of weight loss.

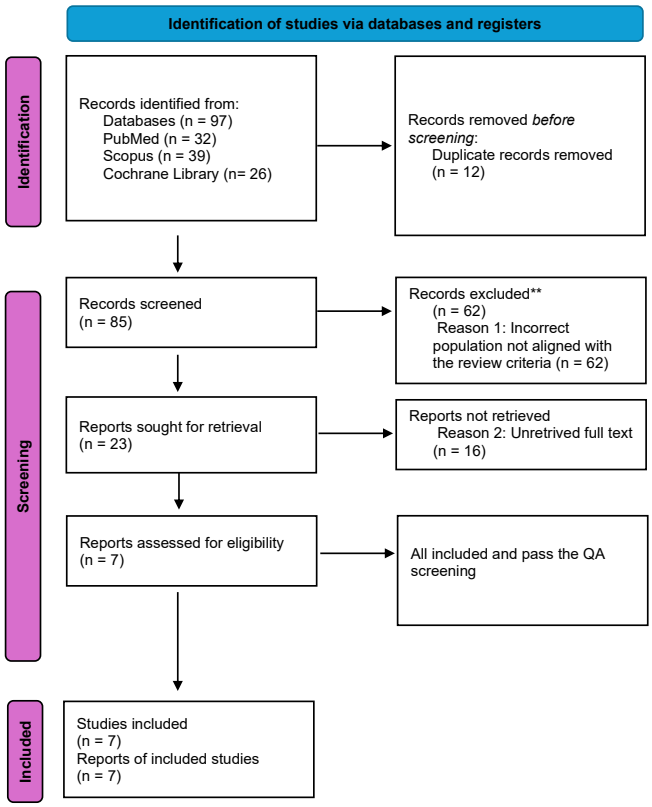


Figure 1: Results of study selection using PRISMA 2020 Flow Diagram (Page et al., 2021)

Safety outcomes were generally reassuring. Reported adverse effects were mild and transient, including hunger, fatigue, nausea, and gastrointestinal discomfort. No studies reported adverse effects on linear growth or pubertal development (Astbury et al., 2019; Evans et al., 2023). Poor adherence was identified as a common challenge, especially in more restrictive interventions such as total diet replacement. Although meal replacement interventions often result in short-term weight loss, evidence on their long-term sustainability remains inconsistent. While Lowe et al. (2018) demonstrated maintenance at 36 months with a home food environment intervention, other studies, such as Xanthopoulos et al. (2013) and Fox et al. (2021), reported significant weight regain within 12 months, particularly when pharmacological or behavioural support was lacking. These findings highlight a persistent gap in sustaining treatment effects, suggesting that current interventions may be insufficient without ongoing management. The variability across studies underscores the need for more rigorous long-term trials to identify effective strategies for maintaining weight loss beyond the initial intervention phase among adolescents.

DISCUSSION

This review evaluated the effectiveness, sustainability, and safety of meal replacement plans (MRPs) in managing obesity among adolescents. The findings demonstrated that MRPs consistently achieved short-term weight reduction within 4 to 12 weeks. These results align with previous adult studies, which also reported clinically significant short-term weight loss through MRPs (Astbury et al., 2019; Khayutin et al., 2023). The effectiveness in the short term can be explained by the substantial calorie restriction imposed by MRPs, which creates a rapid negative energy balance. Additionally, the structured nature of MRPs facilitates dietary decision-making and portion control, resulting in better compliance in the early stages. However, over time, physiological adaptations such as reduced metabolic rate and increased appetite-regulating hormone responses may occur, which can diminish their effectiveness for sustained weight loss. In adolescents, the incorporation of caregiver involvement, digital monitoring tools, and pharmacologic agents further enhanced treatment outcomes, underscoring the importance of multi-component approaches (Berkowitz et al., 2011; Bhagat et al., 2024).

Despite the promising short-term benefits, evidence for long-term sustainability was less consistent. Several studies reported weight regain after 12 months, reflecting the challenges of maintaining adherence once structured interventions ended (Berkowitz et al., 2011; Ferentinou et al., 2023). This echoes findings in adult populations where weight maintenance after MRPs remains a key limitation (Astbury et al., 2019). Adolescents face additional barriers, including psychosocial stressors, social media influence, and family dynamics, which may contribute to poor adherence or relapse (Nagata et al., 2023; Lister et al., 2023). The cost of MRP products may also present a barrier, particularly for families from lower socioeconomic backgrounds, as sustained use could be financially burdensome. This economic factor, combined with limited accessibility in certain communities, further reduces long-term adherence. Interventions with caregiver participation, ongoing behavioural support, and consideration of affordability demonstrated greater potential for sustained weight management (Lambert et al., 2024).

In terms of safety, MRPs were generally well tolerated in adolescents. Reported side effects were mild and transient, including hunger, fatigue, nausea, and gastrointestinal symptoms, with no evidence of adverse effects on growth or pubertal development (Evans et al., 2023; Lee, 2015). However, total diet replacement or highly restrictive regimens were associated with higher

dropout rates and adherence difficulties, suggesting that partial replacement strategies may be more feasible for this age group (Oussaada et al., 2019; Yamada et al., 2023). Current guidelines recommend that MRPs and very-low-energy diets should only be initiated in adolescents with severe obesity, typically defined as BMI ≥35 kg/m² or ≥120% of the 95th percentile, and under close medical supervision (CPG for the Management of Obesity, 2023; Kelly et al., 2013). Furthermore, calorie restriction in this population should generally not fall below 800–1200 kcal/day to prevent nutrient deficiencies, emphasising the importance of individualised planning and monitoring (Andela et al., 2019).

Strengths and Limitations of the Review

The strengths of this review include the use of a systematic methodology guided by PRISMA (Page et al., 2021) and quality appraisal tools (Academy of Nutrition and Dietetics, n.d.), ensuring that findings were synthesized from methodologically sound studies. However, limitations must be acknowledged. The number of eligible studies was small, sample sizes were often limited, and follow-up durations were generally short (Astbury et al., 2019; Berkowitz et al., 2011). Furthermore, heterogeneity in interventions and outcome reporting limited the ability to draw firm conclusions regarding long-term effectiveness (Smyczyńska et al., 2024).

Implications and Future Directions

Future research on MRPs for adolescents should focus on conducting well-designed, large-scale trials with extended follow-up to evaluate the long-term safety and effectiveness of these products. Studies should also address developmental considerations unique to adolescence, including nutritional adequacy for growth and psychological well-being in treatment. Additionally, exploring the cost-effective MRP strategies is necessary to ensure sustainability. Moreover, integrating digital health tools with social support platforms may also enhance adherence, engagement, and sustained behaviour change.

CONCLUSION

The use of MRPs has been demonstrated to be safe and effective in achieving short-term weight loss among adolescents with obesity, with no adverse effects on growth and development. Outcomes were further strengthened when caregivers were involved and when pharmacological or digital health support was integrated. However, the variability in long-term sustainability underscores the need for ongoing support strategies to sustain weight loss and ensure lasting health benefits.

ACKNOWLEDGEMENT

This research project was not funded.

REFERENCES

Academy of Nutrition and Dietetics. (n.d.). *Quality Criteria Checklist: Primary Research*.

Alexander, P. A. (2020). Methodological guidance paper: The art and science of quality systematic reviews. *Review of Educational Research*, *90*(1), 6–23. https://doi.org/10.3102/0034654319854352

Amir-Behghadami, M., & Janati, A. (2020). Population, intervention, comparison, outcomes and study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. *Emergency Medicine Journal*, *37*(6).

Andela, S., Burrows, T. L., Baur, L. A., Coyle, D. H., Collins, C. E., & Gow, M. L. (2019). Efficacy of very low-energy diet programs for weight loss: A systematic review with meta-analysis of intervention studies in children and adolescents with obesity. In Obesity Reviews (Vol. 20, Issue 6, pp. 871–882). Blackwell Publishing Ltd. https://doi.org/10.1111/obr.12830

Astbury, N. M., Piernas, C., Hartmann-Boyce, J., Lapworth, S., Aveyard, P., & Jebb, S. A. (2019). A systematic review and meta-analysis of the effectiveness of meal replacements for weight loss. *Obesity Reviews*, *20*(4), 569–587. https://doi.org/10.1111/obr.12816

Berkowitz, R. I., Wadden, T. A., Gehrman, C. A., Bishop-Gilyard, C. T., Moore, R. H., Womble, L. G., Cronquist, J. L., Trumpikas, N. L., Levitt Katz, L. E., & Xanthopoulos, M. S. (2011). Meal replacements in the treatment of adolescent obesity: A randomized controlled trial. *Obesity*, *19*(6), 1193–1199. https://doi.org/10.1038/oby.2010.288

Bhagat, A., Mehendale, A., & Muneshwar, K. (2024). Study of obesity among adolescents in rural Wardha district. *F1000Research*, *13*, 527. https://doi.org/10.12688/f1000research.148494.1

De Lorenzo, A., Gratteri, S., Gualtieri, P., Cammarano, A., Bertucci, P., & Di Renzo, L. (2019). Why primary obesity is a disease? *Journal of Translational Medicine*, *17*(1). https://doi.org/10.1186/s12967-019-1919-y

Evans, M., de Courcy, J., de Laguiche, E., Faurby, M., Haase, C. L., Matthiessen, K. S., Moore, A., & Pearson-Stuttard, J. (2023). Obesity-related complications, healthcare

resource use and weight loss strategies in six European countries: The RESOURCE survey. *International Journal of Obesity*, *47*(8), 750–757. https://doi.org/10.1038/s41366-023-01325-1

Ferentinou, E., Koutelekos, I., Pappa, D., Manthou, P., & Dafogianni, C. (2023). The impact of the COVID-19 pandemic on childhood obesity: A review. *Cureus*. https://doi.org/10.7759/cureus.45470

Fox, C. K., Kaizer, A. M., Rudser, K. D., Nathan, B. M., Gross, A. C., Sunni, M., Jennifer Abuzzahab, M., Schwartz, B. L., Kumar, S., Petryk, A., Ryder, J. R., & Kelly, A. S. (2016). Meal replacements followed by topiramate for the treatment of adolescent severe obesity: A pilot randomized controlled trial. *Obesity*, *24*(12), 2553–2561. https://doi.org/10.1002/oby.21633

Fox, C. K., Rudser, K., Clark, J., Ryder, J. R., Gross, A. C., Nathan, B. M., Sunni, M., Dengel, D. R., Billington, C. J., Bensignor, M. O., & al., et. (2021). Once-Weekly Exenatide Enhances Weight Loss Maintenance in Adolescents with Severe Obesity: a Randomized, Placebo-Controlled Trial. *Journal of the Endocrine Society*, *5*, A63. https://doi.org/10.1210/jendso/bvab048.128

Gow, M. L., Jebeile, H., House, E. T., Alexander, S., Baur, L. A., Brown, J., Collins, C. E., Cowell, C. T., Day, K., Garnett, S. P., Varady, K. A., & Lister, N. B. (2024). Efficacy, Safety and Acceptability of a Very-Low-Energy Diet in Adolescents with Obesity: A Fast Track to Health Sub-Study. *Nutrients*, *16*(18). https://doi.org/10.3390/nu16183125

Gross, A. C., Freese, R. L., Bensignor, M. O., Bomberg, E. M., Dengel, D. R., Fox, C. K., Rudser, K. D., Ryder, J. R., Bramante, C. T., Raatz, S., Hur, C., & Kelly, A. S. (2024). Financial Incentives and Treatment Outcomes in Adolescents with Severe Obesity: A Randomized Clinical Trial. *JAMA Pediatrics*, *178*(8), 753–762. https://doi.org/10.1001/jamapediatrics.2024.1701

Institute for Public Health (IPH) 2022. (2022). *Technical Report National Health and Morbidity Survey (NHMS) 2022: Adolescent Health Survey, Malaysia*.

Kelly, A. S., Barlow, S. E., Rao, G., Inge, T. H., Hayman, L. L., Steinberger, J., Urbina, E. M., Ewing, L. J., & Daniels, S. R. (2013). Severe obesity in children and adolescents: Identification, associated health risks, and treatment approaches: A scientific statement from the American Heart Association. *Circulation*, *128*(15), 1689–1712. https://doi.org/10.1161/CIR.0B013E3182A5CFB3

Khayutin, S., Kelly, A. S., Fox, C. K., Ryder, J. R., & Gross, A. C. (2023). Opinions from the experts: Experiences of adolescents with severe obesity participating in meal replacement therapy. *Pediatric Obesity*, *18*(2). https://doi.org/10.1111/ijpo.12986

Lambert, D. C., Kane, J., & Newberry, C. (2024). Lifestyle therapy for obesity. *Gastrointestinal Endoscopy Clinics of North America*, *34*(4), 577–589. https://doi.org/10.1016/j.giec.2024.03.003

Lee, G. (2015). Obesity, the epidemic that CAN be stopped? *Journal of Advanced Nursing*, *71*(9), 1971–1972. https://doi.org/10.1111/jan.12584

Lei, S., Scott, C., Lerman, A., Medina-Inojosa, J., Lerman, L., de Andrade, M., Senecal, C., & Lopez-Jimemez, F. (2020). Effectiveness of a weight loss program using digital health in teens and preteens: An observational study of 2,825 adolescents. *Journal of the American College of Cardiology*, *75*(11), 1830. https://doi.org/10.1016/S0735-1097(20)32457-8

Lister, N. B., Baur, L. A., Felix, J. F., Hill, A. J., Marcus, C., Reinehr, T., Summerbell, C., & Wabitsch, M. (2023). Child and adolescent obesity. *Nature Reviews Disease Primers*, *9*(1). https://doi.org/10.1038/s41572-023-00435-4

Mohamed Shaffril, H. A., Samsuddin, S. F., & Abu Samah, A. (2021). The ABC of systematic literature review: The basic methodological guidance for beginners. *Quality & Quantity*, *55*(4), 1319–1346. https://doi.org/10.1007/s11135-020-01059-6

Nagata, J. M., Smith, N., Alsamman, S., Lee, C. M., Dooley, E. E., Kiss, O., Ganson, K. T., Wing, D., Baker, F. C., & Gabriel, K. P. (2023). Association of physical activity and screen time with body mass index among US adolescents. *JAMA Network Open*, *6*(2), e2255466. https://doi.org/10.1001/jamanetworkopen.2022.55466

Oussaada, S. M., van Galen, K. A., Cooiman, M. I., Kleinendorst, L., Hazebroek, E. J., van Haelst, M. M., ter Horst, K. W., & Serlie, M. J. (2019). The pathogenesis of obesity. *Metabolism: Clinical and Experimental*, *92*, 26–36. https://doi.org/10.1016/j.metabol.2018.12.012

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated

guideline for reporting systematic reviews. *International Journal of Surgery*, *88*. https://doi.org/10.1016/j.ijsu.2021.105906

Phan, T.-L. T., Reichard, K. W., & Datto, G. A. (2013). Evaluation of weight loss on a low-calorie meal replacement diet as a potential predictor of weight loss after laparoscopic adjustable gastric banding surgery in adolescents. *Obesity Surgery*, *23*(9), 1384–1388. https://doi.org/10.1007/s11695-013-0935-x

Schünemann, H. J., Higgins, J. P., Vist, G. E., Glasziou, P., Akl, E. A., Skoetz, N., & Guyatt, G. H. (2024). Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In *Cochrane handbook for systematic reviews of interventions* (6th ed., Chap. 14). Cochrane.

Smyczyńska, J., Olejniczak, A., Różycka, P., Chylińska-Frątczak, A., Michalak, A., Smyczyńska, U., Mianowska, B., Pietrzak, I., & Szadkowska, A. (2024). Obesity-related complications including dysglycemia based on 1-h post-load plasma glucose in children and adolescents screened before and after COVID-19 pandemic. *Nutrients*, *16*(15). https://doi.org/10.3390/nu16152568

The Clinical Practice Guidelines (CPG) for the Management of Obesity (2nd ed.). (2023). Malaysia Health Technology Assessment Section (MaHTAS).

World Health Organization. (2024, March 1). *Obesity and overweight*. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Xanthopoulos, M. S., Moore, R. H., Wadden, T. A., Bishop-Gilyard, C. T., Gehrman, C. A., & Berkowitz, R. I. (2013). The association between weight loss in caregivers and adolescents in a treatment trial of adolescents with obesity. *Journal of Pediatric Psychology*, *38*(7), 766–774. https://doi.org/10.1093/jpepsy/jst024

Yamada, T., Kimura-Koyanagi, M., Sakaguchi, K., Ogawa, W., & Tamori, Y. (2023). Obesity and risk for its comorbidities diabetes, hypertension, and dyslipidemia in Japanese individuals aged 65 years. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-29276-7

ORIGINAL ARTICLE

Eating Behaviour and Social Media Addiction among Secondary School Students in Kuantan, Malaysia

Khalisah Adnan[1], Roszanadia Rusali[1,2*], Aliza Haslinda Hamirudin[1], Wan Fathin Fariza Wan Mahmood[1]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Pahang, Malaysia

Keywords:
eating behaviour; social media addiction; adolescents; secondary school students

ABSTRACT

Background: Social media has become integral to adolescents' lives, shaping their eating behaviours through exposure to unregulated content. Visually appealing but energy-dense, a lack of nutrient foods is frequently promoted, influencing adolescents at a formative stage of behaviour development. In Malaysia, the double burden of malnutrition persists alongside unhealthy eating habits. This study aimed to assess eating behaviours and social media addiction among secondary school students in Kuantan, Pahang. **Methods:** A cross-sectional study was conducted among 140 students aged 13 to 15 years at SMK Air Putih, Kuantan, Pahang, using convenience sampling. Data were collected through sociodemographic questions, anthropometric measurements, the Dutch Eating Behaviour Questionnaire (DEBQ), and the Social Media Addiction Scale-Student Form (SMAS-SF). **Results:** Mean scores for eating behaviours were 20.86 ± 8.17 (restrained eating), 22.18 ± 8.56 (emotional eating), and 25.35 ± 7.35 (external eating), with external eating being the most prevalent. Social media addiction levels were classified as non-addicted (15.0 %), low addicted (37.1 %), moderately addicted (45.0 %), and highly addicted (2.9 %). **Conclusion:** External eating was the most common behaviour, followed by emotional and restrained eating. Most students were moderately addicted to social media, while only a small proportion were non-addicted. These findings underscore the need for greater awareness of healthy lifestyles and targeted interventions to mitigate social media addiction while promoting healthy eating habits.

INTRODUCTION

Social media has become deeply embedded in adolescents' lives, particularly after the COVID-19 pandemic, when it served as a primary platform for education and social interaction (Bozzola et al., 2022). Its use is now nearly universal, with platforms such as TikTok, Instagram, and YouTube dominating adolescents' daily engagement (UNICEF, 2020; Bozzola et al., 2022).

Although these platforms offer opportunities for learning and peer connection, it also expose users to unregulated food and lifestyle-related content that may promote both healthy and unhealthy dietary practices (Filippone et al., 2022). Excessive use may lead to addictive behaviours (Andreassen & Pallesen, 2014) and adolescents, being highly impressionable, are especially vulnerable to such influences (UNICEF, 2020).

Adolescence is a critical period for establishing eating habits, which often persist into adulthood and shape long-term health outcomes (Rachmi et al., 2021). Previous studies have shown that social media use is associated with problematic eating behaviours, particularly emotional and external eating (Ayyıldız & Şahin, 2022; Saat et al., 2023) and linked to outcomes such as poor dietary quality, elevated body mass index (BMI), and lower life satisfaction (Seabrook et al., 2016). In Malaysia, the National Health and Morbidity Survey (NHMS, 2024) highlights concerning dietary trends, with only 13% of adolescents consuming sufficient fruits daily, 5% meeting adequate vegetable intake daily, 63% consuming more than one serving of sugar-sweetened beverages daily, and 75% exceeding recommended fat intake. Such unhealthy eating behaviours during adolescence can contribute to both malnutrition and overnutrition, which remain prevalent in Malaysia (NHMS, 2022). The latest survey (NHMS, 2024) reported that 1 in 3 adolescents is overweight or obese, while increasing rates of thinness, overweight, and obesity further highlight the urgent need to address poor dietary practices.

Although existing studies have explored the impact of social media on eating behaviours, most were conducted among university students or focused on broader populations (Aslam et al., 2021; Ayyıldız & Şahin, 2022). In Malaysia, Saat et al. (2023) investigated university students in Klang Valley, but findings may not be

* *Corresponding author.*
E-mail address: roszanadia@iium.edu.my

generalizable to other regions, while little is known about secondary school students in Kuantan, the capital city of Pahang, where an evolving food culture, such as the growing popularity of aesthetic cafés promoted on social media. This may further influence adolescents' dietary behaviours, yet remain unstudied.

Therefore, this study aims to assess the eating behaviours and social media addiction among secondary school students in Kuantan. Findings from this study are expected to provide insights for parents, educators and policymakers in designing interventions to promote digital wellness and healthier eating habits.

MATERIALS AND METHODS

Study Design

This study employed a quantitative cross-sectional design to assess eating behaviours and social media addiction among secondary school students. This study design was selected for its suitability in school-based research, being cost-effective and practical to implement (Wang & Cheng, 2020).

Sampling Method

Study area

The study was conducted at a public secondary school in Kuantan, Pahang. The target population was secondary school students aged 13 to 15 years. The school was randomly selected from eight eligible schools in Kuantan, identified based on the highest student enrolment.

Sample size and sampling

The sample size was calculated using the single proportion formula, using the prevalence of eating disorders among students with potential internet addiction, based on a study conducted by Aslam et al. (2021). The calculated sample size, including a 10% non-response rate, was 152. Convenience sampling was applied to recruit participants who met the inclusion criteria of students aged 13 to 17 years and social media users. Exclusion criteria included hostel students and those diagnosed with eating disorders. Prior to participation, written informed consent was obtained from the parents or guardians of all participants.

Ethical Approval

Ethical approval was obtained from the Kulliyyah Postgraduate and Research Committee (KPGRC) and IIUM Research Ethics Committee (IREC 2025-KAHS/DNS9). Permissions were granted by the Ministry of Education, the Pahang State Education Department and the school principal.

Data Collection

Data collection was conducted during school hours. A questionnaire consisting of four sections was administered to participants by the researcher. The first section collected sociodemographic information, the second assessed participants' eating habits using the Dutch Eating Behavior Questionnaire (DEBQ), the third focused on social media usage and addiction using the Social Media Addiction Scale-Student Form (SMAS-SF) and the fourth comprised anthropometric assessments, which were conducted by the researchers.

Sociodemographic data

The questionnaire includes age, gender, ethnicity, grade level, daily social media use, and preferred platform.

Anthropometric data

The height was measured with the SECA 213 portable stadiometer, the weight was measured with the Omron BIA digital scale, and BMI was classified using WHO BMI-for-age growth charts (5-19 years) (WHO, 2007).

Dutch Eating Behaviour Questionnaire (DEBQ)

A Malay version of a 33-item questionnaire assessing restrained (10 items), emotional (13 items), and external eating (10 items) behaviours on a 5-point Likert scale (van Strien et al., 1986). Higher scores indicate stronger eating behaviour tendencies. Reliability analysis showed Cronbach's alpha values of 0.897, indicating acceptable internal consistency (Tavakol & Dennick, 2011).

Social media addiction scale-student form (SMAS-SF)

The Malay version questionnaire consists of 29 items assessing four domains (virtual tolerance, communication, problems, and information) on a 5-point Likert scale (Sahin, 2018). Total scores classify respondents as non-addicted (29 to 51), low addicted (52 to 74), moderate (75 to 97), or highly addicted (98 to 120). Reliability analysis showed Cronbach's alpha values of 0.805, indicating acceptable internal consistency (Tavakol & Dennick, 2011).

Statistical Analysis

Data was analysed using SPSS Version 20. Descriptive statistics (frequencies, percentages, means, standard deviations) were used to summarise sociodemographic,

anthropometric, eating behaviour, and social media addiction data. Fisher's exact test was used to assess associations between categorical variables. An independent t-test compared mean DEBQ scores between genders. Statistical significance was set at $p < 0.05$.

RESULTS

Sociodemographic Characteristics of Respondents

Table 1 presents an overview of the sociodemographic characteristics of the respondents, including gender, age, ethnicity, and daily social media usage. Data are presented in frequency (n) and percentage (%), analysed using descriptive analysis to provide a clear understanding of the distribution within the sample population. Among 140 respondents, 36.4% (n = 51) were male, and 63.6% (n = 89) were female, indicating a higher proportion of female participants. Most respondents were 14 years old (39.3%, n = 55), followed by 13 years old (36.4%, n = 51), with the smallest group being 15 years old (24.3%, n = 34).

In terms of ethnicity, the majority of respondents were Chinese (81.4%, n = 114), followed by Malay (11.4%, n = 16), Indian (5%, n = 7), and others (2.1%, n = 3), including Iban. Regarding daily social media usage, 42.9% (n = 60) of respondents spent ≤2 hours per day on social media, whereas 57.1% (n = 80) spent >2 hours per day, indicating that a majority engage in more than two hours of social media use daily.

Table 1: Demographic characteristics of the respondents (n = 140)

Characteristics	Frequency (n)	Percentage (%)
Gender		
Male	51	36.4
Female	89	63.6
Age		
13	51	36.4
14	55	39.3
15	34	24.3
Ethnic		
Chinese	114	81.4
Indian	7	5.0
Malay	16	11.4
Others	3	2.1
Time spent on social media each day		
≤ 2 hours	60	42.9
> 2 hours	80	57.1

Anthropometric Findings

Table 2 shows anthropometric measurements, including height, weight, BMI values and BMI classifications for 140 respondents. The mean height was 156.58 cm (SD = 7.5 cm), the mean weight was 48.47 kg (SD = 11.8 kg) and the mean BMI was 19.78 $\mathrm{kg/m^2}$ (SD = 4.1 $\mathrm{kg/m^2}$).

Regarding BMI classification, most respondents (75%, n = 105) had normal BMI. Overweight respondents accounted for 10% (n = 14), while 7.1% (n = 10) were classified as obese. Thinness was observed in 4.3% (n = 6), while severe thinness was 3.6% (n = 5). This indicates that although the majority of students had a normal BMI, a notable proportion were either underweight or overweight.

Table 2: Anthropometric data of the respondents (n = 140)

Variable	Mean (SD)	n (%)
Height (cm)	156.58 (7.5)	
Weight (kg)	48.47 (11.8)	
BMI (kg/m^2)	19.78 (4.1)	
BMI classification		
Severe thinness		5 (3.6)
Thinness		6 (4.3)
Normal		105 (75.0)
Overweight		14 (10.0)
Obese		10 (7.1)

Table 3 presents BMI classification by gender, analysed using Fisher's exact test. Among males, severe thinness was 2.0% (n = 1), thinness 3.9% (n = 2), normal 78.4% (n = 40), overweight 7.8% (n = 4), and obese 7.8% (n = 4). Among females, severe thinness was 4.5% (n = 4), thinness 4.5% (n = 4), normal 73.0% (n = 65), overweight 11.2% (n = 10), and obese 6.7% (n = 6). No significant association between gender and BMI classification was observed (p = 0.932).

Table 3: BMI classification according to gender (Fisher's exact test, n = 140)

Variables	Male (n = 51)		Female (n = 89)		Total (n = 140)		*p*-value
	n	%	n	%	n	%	
BMI classification							
Severe thinness	1	2.0	4	4.5	5	3.6	0.932
Thinness	2	3.9	4	4.5	6	4.3	
Normal	40	78.4	65	73.0	105	75.0	
Overweight	4	7.8	10	11.2	14	10.0	
Obese	4	7.8	6	6.7	10	7.1	

Eating Behaviour

Table 4 shows descriptive statistics and gender comparison of eating behaviour scores measured using the DEBQ. The mean score for restrained eating was 20.86 (SD = 8.17), indicating a lower tendency to intentionally restrict food intake for weight control, while emotional eating scored 22.18 (SD = 8.56), suggesting a tendency to eat in response to emotions such as stress. External eating had the highest mean score at 25.35 (SD = 7.35), showing that respondents were most influenced by environmental food cues such as smell.

Independent t-test analysis revealed no significant differences in restrained, emotional, or external eating behaviours between male and female respondents. Although females reported slightly higher in restrained eating and males demonstrated higher emotional and external eating, these differences were not statistically significant.

Table 4: Eating Behaviour scores across DEBQ domain according to gender (Independent t-test, n = 140)

Variable	Total		Male (n = 51)		Female (n = 89)		Mean differences (95% CI)	*p*-value
	Mean	SD	Mean	SD	Mean	SD		
Restrained eating	20.86	8.17	20.43	7.93	21.10	8.34	-0.67 (-3.52, 2.18)	0.64
Emotional eating	22.18	8.56	23.71	9.99	21.30	7.54	2.40 (-0.80, 5.61)	0.14
External eating	25.35	7.35	26.92	7.03	24.45	7.42	2.47 (-0.06, 5.00)	0.06

Social Media Addiction

Figure 1 presents social media addiction levels based on SMAS-SF scores. Among respondents, 15% (n = 21) were non-addicted, 37.1% (n = 52) low addicted, 45% (n = 63) moderately addicted, and 2.9% (n = 4) highly addicted.

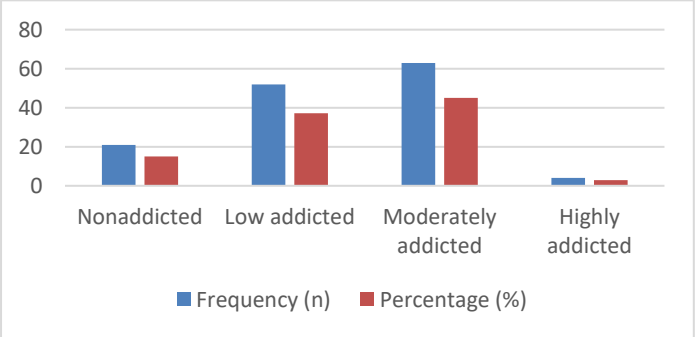


Figure 1: Distribution of Social Media Addiction Category (n=140)

Table 5 shows the distribution of social media addiction categories according to ethnicity, BMI classification, and gender using Fisher's exact test. Ethnicity showed a significant association with addiction levels ($p = 0.02$), with Chinese respondents representing the majority in most categories. Normal BMI was the most common across all addiction categories, with no significant association observed ($p = 0.40$). Females were the majority across all addiction levels, while males had smaller proportions, and there were no significant associations with gender ($p = 0.25$).

Table 5: Social Media Addiction according to ethnicity, BMI, and gender (n = 140)

Variables	Non addicted (n = 21)		Low addicted (n = 52)		Moderately addicted (n = 63)		Highly addicted (n = 4)		*p*-value
	n	%	n	%	n	%	n	%	
Ethnic									
Chinese	18	85.7	42	42.3	52	82.5	2	50.0	
Indian	2	9.5	4	1.9	1	1.6	0	0.0	0.02*
Malay	1	4.8	5	53.8	10	15.9	0	0.0	
Others	0	0.0	1	1.9	0	0.0	2	50.0	
BMI classification									
Severe thinness	2	9.5	2	1.9	1	1.6	0	0.0	
Thinness									0.40
Normal	1	4.8	1	1.9	4	6.3	0	0.0	
Overweight	17	81.0	40	76.9	46	73.0	2	50.0	
Obese	0	0.0	6	11.5	7	11.1	1	25.0	
	1	4.8	3	5.8	5	7.9	1	25.0	
Gender									
Male	5	23.8	20	38.5	26	41.3	0	0.0	0.25
Female	16	76.2	32	61.5	37	58.7	4	100.0	

* $p < 0.05$, Fisher's exact test

DISCUSSION

Eating Behaviour

The sociodemographic profile showed that 75% of participants had normal BMI, 10% were overweight, 7.1% obese, 4.3% thin, and 3.6% severely thin, reflecting Malaysia's double burden of malnutrition (NHMS, 2022). The high prevalence of social media use, with 57.1% spending more than 2 hours daily, suggests that digital platforms are exposing adolescents to food-related content, such as influencer posts or fast-food marketing (Filippone et al., 2022), may increase external and emotional eating tendencies. The female-dominated sample may further influence emotional eating, as prior studies highlight female adolescents' vulnerability to emotional and external triggers (Saat et al., 2023).

Although the study initially posited that emotional and external eating would be more dominant than restrained eating, the observed pattern ultimately reflected a stronger influence of external cues on eating behaviour. This aligns with Ayyıldız & Şahin (2022), who found that individuals with social media addiction are more prone to external eating. External eating is linked to obesity, as overweight and obese individuals are sensitive to environmental cues that encourage overeating, such as variety and easy access to food (Benbaibeche et al., 2023). The elevated mean for external eating suggests that students are highly responsive to environmental food cues, such as smell, appearance, and accessibility (Stroebele & De Castro, 2004). Emotional eating showed moderate prevalence, indicating eating in response to emotional states like stress, sadness, or boredom, which is consistent with adolescent psychological development, where emotional regulation is still maturing (Best & Ban, 2021). This pattern is also in line with Ayyıldız & Şahin (2022), who noted that emotional eating tends to increase among those with higher social media engagement.

Restrained eating was the least prevalent, which suggests that conscious food restriction for weight control is less frequent, possibly because adolescents prioritise immediate gratification over long-term dietary control (Linardon, 2018). Social media exposure, which frequently features appetising food content, may weaken self-control and promote impulsive eating.

Regarding gender, contrary to Saat et al. (2023), males reported higher emotional and external eating scores, while females scored higher on restrained eating. This aligns with Dane & Bhatia (2023), who suggested that females engage more in dieting due to thin body ideals, social comparison, and appearance-focused social media, whereas males often pursue muscular ideals and may engage in emotional and external eating in response to

conflicting body expectations. These gender differences highlight how social media affects males and females differently, highlighting the importance of addressing gender-specific media influences and promoting social media literacy to prevent unhealthy eating habits.

Social Media Addiction

The overall pattern indicates that moderate social media addiction is more prevalent than high addiction among adolescents, suggesting that the initial hypothesis of high addiction was not fully supported. This observation is consistent with global trends showing that adolescents are active social media users and typically maintain at least one active account (UNICEF, 2020).

The predominance of moderate addiction reflects the immersive design of platforms such as TikTok, Instagram, and YouTube, which use algorithms and influencer content to sustain and drive engagement (Yusoff et al., 2023). The small proportion of highly addicted students suggests severe addiction is uncommon, possibly due to school regulations, parental monitoring, and cultural attitudes towards technology, particularly among Chinese students who dominated the sample. RedNote and Douyin were the most used platforms among this population. The predominantly female sample may also influence addiction patterns, as gender-specific social pressures shape engagement (Saat et al., 2023).

No significant differences were found in addiction categories when compared by gender or BMI classification, with normal BMI predominating across all groups. This contrasts with Saat et al. (2023) who reported female adolescents as more susceptible to excessive screen time and increased emotional and external eating behaviours. Also with Ayyıldız & Şahin (2022) who linked social media addiction to poor nutritional status. In this study, adolescents' growth phase, higher physical activity levels due to school routines and extracurricular activities, or naturally greater energy expenditure, may have helped maintain normal BMI despite varying levels of social media use (Silva et al., 2022). However, barriers such as lack of time, low motivation, and limited access to facilities suggest that adolescents remain at risk of reduced activity over time, underscoring the importance of early intervention.

In contrast, ethnicity showed a significant association with social media addiction categories, with Chinese students dominating all levels, while Malay students were more prevalent in the moderately addicted group. These findings are consistent with Cummings (2014) who emphasised the influence of cultural factors on social media engagement.

CONCLUSION

The findings of this study indicate that external eating is the most prevalent eating behaviour among secondary school students, followed by emotional eating, with restrained eating being the least common. This suggests that adolescents are highly sensitive to external food cues, which may be influenced by exposure to food-related content on social media platforms. Regarding social media addiction, the majority of students were moderately addicted, reflecting global trends of high engagement among adolescents with digital platforms. These findings provide valuable insights for educators, parents, and policymakers in supporting adolescent well-being. Strategies may include school-based nutrition programs, digital literacy campaigns, and health promotion initiatives, while collaboration between the Ministries of Health and Education could help regulate food-related content on social media to encourage healthier lifestyles.

This study had several limitations, including a predominantly Chinese sample that limits generalizability, only students aged 13 to 15 were included, which reduces the representativeness, and potential misunderstandings from the Malay-language questionnaire, although clarifications were provided in English. Despite these challenges, the strengths of this study include direct anthropometric measurement by the researcher, which minimised self-report bias and ensured accurate BMI classification. The majority Chinese sample also offers valuable insight into eating behaviour patterns which are often underrepresented in local studies. For future research, involving more diverse samples, provide bilingual surveys, and consider other factors such as physical activity, sleep, stress, and socioeconomic status. Multi-school studies with larger participation would further improve generalizability and provide deeper insights.

ACKNOWLEDGEMENT

Our sincere gratitude to all institutions and individuals involved for the support and encouragement throughout this study.

REFERENCES

Andreassen, C. S., & Pallesen, S. (2014). Social network site addiction: An overview. *Current Pharmaceutical Design, 20*(25), 4053–4061. https://doi.org/10.2174/1381612811319990616

Aslam, M., Khan, J. G., Khan, G., Asif, H., Atta, N., Rafique, A., Tahir, T., Ahmed, N., & A., A. (2021). Impact of

social media and peer pressure on eating behaviors of adolescents. *Pakistan BioMedical Journal, 4*(2). https://doi.org/10.54393/pbmj.v4i2.147

Ayyıldız, F., & Şahin, G. (2022). Effect of social media addiction on eating behavior, body weight, and life satisfaction during pandemic period. *British Food Journal, 124*(9). https://doi.org/10.1108/BFJ-01-2022-0018

Benbaibeche, H., Saidi, H., Bounihi, A., & Koceir, E. A. (2023). Emotional and external eating styles associated with obesity. *Journal of Eating Disorders, 11*(1), 1–7. https://doi.org/10.1186/s40337-023-00797-w

Best, O., & Ban, S. (2021). Adolescence: Physical changes and neurological development. *British Journal of Nursing, 30*(5), 272–275. https://doi.org/10.12968/bjon.2021.30.5.272

Bozzola, E., Spina, G., Agostiniani, R., Barni, S., Russo, R., Scarpato, E., Di Mauro, A., Di Stefano, A. V., Caruso, C., Corsello, G., & Staiano, A. (2022). The use of social media in children and adolescents: Scoping review on the potential risks. *International Journal of Environmental Research and Public Health, 19*(16), 9960. https://doi.org/10.3390/ijerph19169960

Cummings, J. R. (2014). Contextual socioeconomic status and mental health counseling use among U.S. adolescents with depression. *Journal of Youth and Adolescence, 43*(7), 1151–1162. https://doi.org/10.1007/s10964-013-0021-7

Dane, A., & Bhatia, K. (2023). The social media diet: A scoping review to investigate the association between social media, body image, and eating disorders amongst young people. *PLOS Global Public Health, 3*(3), e0001091. https://doi.org/10.1371/journal.pgph.0001091

Filippone, L., Shankland, R., & Hallez, Q. (2022). The relationships between social media exposure, food craving, cognitive impulsivity, and cognitive restraint. *Journal of Eating Disorders, 10*(1), 1–12. https://doi.org/10.1186/s40337-022-00698-4

Linardon, J. (2018). The relationship between dietary restraint and binge eating: Examining eating-related self-efficacy as a moderator. *Appetite, 127*, 126–129. https://doi.org/10.1016/j.appet.2018.04.026

Ministry of Health Malaysia. (2022). *Adolescent Health Survey 2022*. Institute for Public Health. https://iku.gov.my/images/nhms-2022/Report_Malaysia_nhms_ahs_2022.pdf

Ministry of Health Malaysia. (2024). *National Health and Morbidity Survey (NHMS) 2024*. Institute for Public Health. https://iku.nih.gov.my/NHMS-2024

Rachmi, C. N., Jusril, H., Ariawan, I., Beal, T., & Sutrisna, A. (2021). Eating behaviour of Indonesian adolescents: A systematic review of the literature. *Public Health Nutrition, 24*, S84–S97. https://doi.org/10.1017/S1368980020002876

Saat, N. Z. M., Hanawi, S. A., Chew, N. H. H., Ahmad, M., Farah, N. M. F., Kadar, M., Yahya, H. M., Warif, N. M. A., & Daud, M. K. M. (2023). The association of eating behaviour with physical activity and screen time among adolescents in the Klang Valley, Malaysia: A cross-sectional study. *Healthcare, 11*(9), 1260. https://doi.org/10.3390/healthcare11091260

Sahin, C. (2018). Social Media Addiction Scale–Student Form: The reliability and validity study. *Turkish Online Journal of Educational Technology, 17*(1), 169–182.

Seabrook, E. M., Kern, M. L., & Rickard, N. S. (2016). Social networking sites, depression, and anxiety: A systematic review. *JMIR Mental Health, 3*(4). https://doi.org/10.2196/mental.5842

Silva, R. M. F., Mendonça, C. R., Azevedo, V. D., Memon, A. R., Silva Noll, P. R. E., & Noll, M. (2022). Barriers to high school and university students' physical activity: A systematic review. *PLOS ONE, 17*(4), e0265913. https://doi.org/10.1371/journal.pone.0265913

Stroebele, N., & De Castro, J. M. (2004). Effect of ambience on food intake and food choice. *Nutrition, 20*(9), 821–838. https://doi.org/10.1016/j.nut.2004.05.012

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

UNICEF. (2020). *Use of social media by children and adolescents in East Asia: Opportunities, risks and harms*. https://www.unicef.org/malaysia/media/1506/file/

Van Strien, T., Frijters, J. E. R., Bergers, G. P. A., & Defares, P. B. (1986). The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *International Journal of Eating Disorders, 5*(2), 295–315. https://doi.org/10.1002/1098-108x(198602)5:2

Wang, X., & Cheng, Z. (2020). Cross-sectional studies: Strengths, weaknesses, and recommendations. *Chest, 158*(1), S65–S71. https://doi.org/10.1016/j.chest.2020.03.012

WHO. (2007). *BMI-for-age (5–19 years)*. World Health Organization. https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age

Yusoff, H. A. M., Hamzah, M. R., Manaf, A. R. A., Suanda, J., Mutalib, S. A., & Ayub, S. H. (2023). Health-related social media use among youth in Malaysia. *AIP Conference Proceedings, 2544*(1). https://doi.org/10.1063/5.0132823

ORIGINAL ARTICLE

Knowledge, Attitudes, and Practices of Laboratory Biosafety Among Undergraduate Students in a Malaysian Public University: Lab Biosafety KAP among Malaysia Undergraduates

Syaza Nasyitah Shamsol Kamal[1], Norfazlina Shaharuddin[1,2], Hanani Ahmad Yusof @ Hanafi[1]*

[1]Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia
[2]Department of Biology, Centre for Foundation Studies, International Islamic University Malaysia, 26300 Gambang, Pahang, Malaysia

Keywords:
laboratory biosafety; knowledge, attitude and practice; undergraduate students; health science program

ABSTRACT

Background: Laboratory biosafety is comprised of a crucial set of rules and regulations to be followed by the laboratory users. However, students' awareness of the biosafety regulations and practices is unknown since there is no current data on laboratory biosafety knowledge. Therefore, this study aims to evaluate the knowledge, attitude and practice (KAP) levels regarding laboratory biosafety among health sciences undergraduates and determine the factors influencing their knowledge. **Methods:** A modified and adapted questionnaire was administered to the voluntary respondent who fulfill the inclusion and exclusion criteria. The questionnaires which consist of 31 questions related to the knowledge, attitude and practice on laboratory biosafety were distributed through online platform and face to face approaches. Spearman's test was used to determine the correlation between the KAP elements, while independent t-tests was used to assess the association between the knowledge and demographic background of the respondent including the gender and briefing status on laboratory biosafety. One way ANOVA and Kruskal Wallis was applied to compare the differences of biosafety knowledge across academic years and faculties, respectively. **Results:** A total of 388 students voluntarily participated in the study. The survey showed that the students have a moderate knowledge on laboratory biosafety whilst have a good attitude and practice towards it. While no significant correlation was found between knowledge and attitude ($r=0.046$, $p=0.371$), practice showed a moderate positive correlation with attitude ($r=0.417$, $p<0.001$) and weak positive correlation with knowledge ($r=0.135$, $p=0.008$). Knowledge levels were significantly associated with the year of study ($p<0.001$) and briefing status on laboratory biosafety ($p=0.013$). However, knowledge score did not differ significantly by gender or faculty. **Conclusion:** Undergraduate health science students exhibit moderate knowledge on laboratory biosafety, despite demonstrating good perceptions and adherence to biosafety practices. The observed knowledge gaps may be attributed to insufficient briefing on the subject. Improving laboratory biosafety education could enhance student's preparedness for laboratory work throughout their academic programme.

INTRODUCTION

Biosafety refers to the integrated implementation of laboratory practices, protocols and standards, combined with laboratory facilities and safety equipment when working with potentially hazardous biological materials including infectious agents such as bacteria, parasites, fungi, virus and prion (Bayot & Limaiem, 2019; World Health Organization, 2023). Strict compliance with biosafety protocols is essential to prevent laboratory accidents, particularly laboratory-acquired infections (LAIs) – any symptomatic or asymptomatic disease contracted through laboratory work or related activities (Peng et al., 2018; Gugnani & Randhawa, 2020). The LAIs cases have been documented since 1885 and have been reported in various countries. A noble incident in Malaysia involved exposure to Brucella sp., affecting over 50 laboratory workers, with several experiencing adverse health effects. The investigation revealed that primary contributing factors included failure to comply with standard operating procedures (SOPs), absence of a directional airflow system in the laboratory, and improper use of biosafety cabinet (Sam et al., 2012). A parallel case in Singapore demonstrated similar biosafety failures, where insufficient training on Class III biosafety cabinet led to occupational exposure to SARS-associated coronavirus (SARS-CoV) (Lim et al., 2004).

Effective training plays a crucial role in enhancing laboratory biosafety awareness among users, particularly health science students who frequently engage in biological laboratory work. However, biosafety knowledge

* *Corresponding author.*
E-mail address: hanani@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

levels vary among these students. A report revealed that most students are familiar with emergency protocols for incidents involving fire, chemical spills and gas leaks, but awareness appears notably lower regarding specific biosafety cabinet procedures and proper Personnel Protective Equipment (PPE) application (Walters et al., 2017).

Laboratory incidents have been numerously reported worldwide, some of the cases cause serious injuries, even death (Chen et al., 2023). Incorrect use of chemicals, lack of safety equipment, failure to apply PPE completely are among the factors caused the incidence, which ultimately caused by poor compliance towards safety protocol set for the laboratory. Health science programmes including medicine, dentistry, pharmacy, and allied health sciences highly offer laboratory-based education involving potentially hazardous materials such as biochemistry, haematology, histology, microbiology, parasitology and any biological laboratory practical. Frequent exposure to these hazardous and infectious materials increases the likelihood of health risk for the student, particularly if the safety measures are not adhere up to optimal level.

Although a number of awareness surveys have been conducted at numerous institutions and colleges globally, data remained limited for Malaysian public university undergraduates. This study evaluated laboratory biosafety awareness among health science undergraduates at International Islamic University Malaysia (IIUM). In specific, this study intended to evaluate the level of knowledge, attitude and practice on laboratory biosafety among the student, identify key factors influencing their level of knowledge including gender, faculty, academic year and prior laboratory biosafety training, and analyse correlations between knowledge, attitudes and practices on laboratory biosafety.

The findings provide valuable insight for enhancing biosafety protocols in academic institution offering laboratory-based education, ultimately contributing to improved safety standards for undergraduate laboratory training.

MATERIALS AND METHODS

Ethical Approval

Ethical approval for the survey was granted from IIUM Research Ethics Committee (IREC) (Approval ID: IREC 2024 (UG)-046.

Study Design, Area and Participants

This research employed quantitative cross-sectional approach, involving undergraduate students from all faculties at the IIUM Health Sciences campus.

Sampling

The study applied convenience sampling, targeting eligible undergraduate students at the IIUM Health Sciences campus who had experience handling biological samples or materials during laboratory activities. Postgraduate students and those who worked exclusively with non-biological samples were excluded. The participation was entirely on voluntary basis.

A total of 388 undergraduate students from all faculties were recruited. The sample size was estimated using the Krejcie and Morgan table (Krejcie & Morgan, 1970), estimating 381 respondents with additional 10% non-response rate. This calculation was based on the total population of 3682 registered undergraduate students at the IIUM Health Sciences campus during that academic year.

Questionnaires Development and Distribution

The questionnaire was adapted with minor modifications from six published studies (Burdass et al., 2005; World Health Organization, 2010; El-Gilany et al., 2017; Meechan et al., 2020; Mehta & Diwakar, 2021; Ahmed et al., 2022). The questionnaire underwent rigorous validation, reliability and pilot testing prior to implementation. The questionnaires comprised five sections; Section 1: Informed Consent Documentation, Section 2: Demographic Information (including academic programme and year of study), Section 3: Knowledge Assessment on Laboratory Biosafety, Section 4: Attitude Assessment on Laboratory Biosafety, Section 5: Practice Assessment on Laboratory Biosafety. For Section 3, each correct response (selected from four options) received one mark, while incorrect answers or “I do not know” response were scored zero. Assessment for Section 4 was measured using a 5-point Likert scale ranging (1 point = Strongly Disagree to 5 points = Strongly Agree). Similarly, Section 5 was evaluated using 5-points Likert scale with responses scored as follows: 1 point for ‘never’, 2 points for ‘rarely’, 3 points for ‘occasionally’, 4 points for ‘frequently’ and 5 points for ‘always’. The classification of knowledge, attitude and practice level was established according to the framework developed from Mehta and Diwakar (2021) as presented in Table 1.

Table 1: Classification of score for level of knowledge, attitude and practice

Level of KAP*	Score		
	Knowledge	Attitude	Practice
Good	8 - 10	38 - 50	42 - 55
Moderate	5 - 7	25 - 37	27 - 41
Poor	0 - 4	0 - 24	0 - 26

**Good level was determined by the ≥75% scores obtained in each category. Moderate level was determined by the 50 to 74% scores obtained in each category. Poor level was determined by the <50% scores obtained in each category.*

The modified questionnaires underwent expert validation and reliability of instruments by two laboratory biosafety specialists, from different faculties within the host institution. Following validation and reliability test (Cronbach's alpha = 0.873), the instrument was administered in a pilot study involving 38 volunteer undergraduate students from the IIUM Health Campus.

The questionnaires were administered both offline and online from April 2024 to May 2024. For offline participation, respondents received printed questionnaire, while the online participants accessed the survey form through Google Form link distributed through Whatsapp or Telegram platform. All respondents were gently reminded that participation was on voluntarily basis and that they were under no-obligation to respond. All collected data were treated in a confidential manner.

Data Analysis

All responses were compiled and statistically analysed using IBM Statistical Package for the Social Sciences (SPSS) software (version 20.0). Normality test was run prior to statistical testing. Descriptive statistics were employed to evaluate respondents' level of knowledge, attitude and practice regarding laboratory biosafety, while correlation analysis examined relationships between these variables.

To determine potential influencing factors (including gender, academic programme, faculty affiliation, and laboratory briefing experience) on biosafety knowledge levels, statistical tests were employed including independent t-test, one-way ANOVA and Kruskal-Wallis tests. The p-value <0.05 is deemed as significant value for all statistical tests.

RESULT

Demographic Data

The study recruited 388 undergraduate students enrolled in various programmes across six faculties; Science (KOS), Allied Health Sciences (KAHS), Pharmacy (KOP), Dentistry (KOD), Medicine (KOM) and Nursing (KON). The majority of respondents were from KAHS (243 students, 62.6%), with a predominant representation of female students (343, 88.4%). Participants spanned all academic years (Year 1 through year 4/ final year), with the highest proportion from Year 2. Approximately one-third of respondents reported having received laboratory biosafety briefing, most of whom were briefed during their first year of study (Table 2).

Table 2: Distribution of students by gender, faculty, year of study, receipt of briefing and the frequency of briefing received

Variable		Frequency, N	Percentage (%)
Gender			
	Male	45	11.6
	Female	**343**	**88.4**
Kulliyyah			
	Allied Health Science	**243**	**62.6**
	Dentistry	13	3.4
	Medicine	36	9.3
	Nursing	44	11.3
	Pharmacy	18	4.6
	Science	34	8.8
Year of study			
	First year	80	20.6
	Second year	**129**	**33.2**
	Third year	79	20.4
	Final year	100	25.8
Have you ever been briefed about biosafety?			
	Yes	**297**	**76.5**
	No	91	23.5
If yes, how often were these briefings conducted?			
	Every year	52	17.1
	First year of study	**181**	**59.5**
	Other year of study	71	23.4

**The highest frequency and percentage were bold*

Level of Knowledge, Attitude and Practice on Laboratory Biosafety

The study found that the majority of respondents (n = 202, 52.1%) demonstrated moderate knowledge on laboratory biosafety, with a mean score of 7.27± 1.43. However, they exhibited good attitudes (n = 378, 97.4%; mean score = 46.84 ± 6.2) and good practices (n = 304, 78.4%; mean score 45.96 ± 7.2). All scores are reported as mean ± standard deviation (Table 3).

Table 3: Level and mean score of knowledge, attitude and practice on laboratory biosafety

Category	Frequency	Percentage (%)	Mean Score	Standard deviation	Interpretation
Knowledge					
Good	173	44.6	7.27	1.427	Moderate knowledge
Moderate	**202**	**52.1**			
Poor	13	3.4			
Attitude					
Good	**378**	**97.4**	46.84	6.202	Good attitude
Moderate	2	0.5			
Poor	8	2.1			
Practice					
Good	**304**	**78.4**	45.96	7.197	Good practice
Moderate	80	20.6			
Poor	4	1.0			

**The highest frequency and percentage were in bold*

Knowledge of laboratory biosafety

Ten questions were prepared to assess the students' knowledge of laboratory biosafety, covering key elements such as hand hygiene practices, waste management, emergency equipment handling, PPE usage and biosafety cabinet utilisation. Almost all respondents answered the question regarding the function of wearing an eye shield correctly (n = 384, 100%). This was followed by questions related to biohazard waste disposal bin (n = 368, 94.8%) and source of infection in a laboratory (n = 353, 91%).

Table 4: Students' knowledge of laboratory biosafety

Items	Percentage of correct response to question, n (%)
How many steps of hand washing do you know?	315 (81.3%)
What type of fire extinguisher is the most suitable in the laboratory?	256 (66.0%)
What is the suitable percentage of alcohol should have in alcohol-based hand sanitizers?	254 (65.5%)
Biohazard waste should be disposed in what type of bin?	368 (94.8%)
Which of the following is the most common source of infection in a laboratory?	353 (91.0)
Which Personal Protective Equipment (PPE) are recommended for handling pathogenic microorganisms?	240 (61.9%)
What is the primary purpose of biosafety cabinet?	334 (86.2%)
What are the correct steps to clean a biological spill?	128 (33.0%)
Which glove is appropriate to use when working with potentially contagious materials?	188 (48.5%)
What is the purpose of wearing eye shield?	384 (99%)

The lowest scores were recorded for questions related to the type of gloves to use when handling potentially contagious materials (n = 188, 48.5%) and the correct procedure for cleaning a biological spill (n = 128, 33%) (Table 4). In overall, the data indicated strong student knowledge of eye shield importance but identified significant gaps in understanding the correct biological spill cleanup procedure.

Attitude towards laboratory biosafety

Ten statements assessing attitudes toward laboratory biosafety were employed, covering key areas including hand hygiene practices, waste management, emergency equipment handling, PPE use and biosafety cabinet utilisation. Results showed that the 'Strongly Agree' response was most frequently selected by the students. This was particularly evident for statements emphasising the importance of hand hygiene procedures (n = 334, 86.1%), and biological waste management (n = 334, 86.1%) (Table 5) This demonstrated a notably strong positive attitude, particularly toward hand hygiene and biological waste management.

Practice of laboratory biosafety

The students' practice of laboratory biosafety was assessed using eleven statements representing key areas of hand hygiene practices, waste management, emergency equipment handling, PPE use and biosafety cabinet utilisation. The most consistently reported practice was handwashing after handling potentially infectious materials, with 82.2% (n = 319) of students selecting the 'always' response. This was followed by proper disposal practice of biological waste in designated bins and PPE use during laboratory work (Table 6). However, the majority of students reported never using the fire extinguisher in the laboratory. Overall, the results indicated a strong

commitment to core daily safety practices but a lack of engagement with emergency equipment.

Table 5: Students' attitude towards laboratory biosafety

Items	Strongly agree
It is important to follow proper hand hygiene procedures at all times.	334 (86.1%)
Everyone should know how to use fire extinguisher	268 (69.1%)
Alcohol-based hand rub or hand sanitizer should be used every time finish lab work	284 (73.2%)
Using proper methods to handle biological waste is crucial.	334 (86.1%)
I believe a crowded workspace promotes the spread of infections.	265 (68.3%)
I believe wearing PPE could protect me from laboratory hazard	275 (70.9%)
A biosafety cabinet is important for ensuring safety when working with pathogens.	288 (74.2%)
Infectious diseases can be reduced by adhering to proper protocols when managing biological spills.	309 (79.6%)
I believe that wearing gloves is essential for my personal protection when working with germs.	313 (80.7%)
Wearing eye protection is important for my safety when handling pathogens that can cause splashes or sprays.	323 (83.2%)

Correlation between knowledge, attitude and practice on laboratory biosafety

Based on the interpretation of correlation value from Schober et al. (2018) (Table 7), there was negligible correlation observed between knowledge and attitude on laboratory biosafety (r=0.046, n=388, p=0.371). Meanwhile, a weak positive correlation was observed between knowledge and practice (r=0.135, n=388, p=0.008) and attitude and practice (r=0.417, p<0.001) (Table 8). These results indicated that a positive attitude correlated most strongly with safe practices, whereas knowledge had a minimal direct impact on both attitudes and behavior.

Table 7: Interpretation of the correlation value

Correlation value	Interpretation
0.00-0.10	Negligible correlation
0.10-0.39	Weak correlation
0.40-0.69	Moderate correlation
0.70-0.89	Strong correlation
0.90-1.00	Very strong correlaion

Note. The range was adopted from Schober and Schwarte.

Table 6: Students' practice on laboratory biosafety (n=388)

Items	Always
I wash my hands with soap and water after wearing glove.	222 (57.2%)
I use a fire extinguisher in the laboratory according to the specified guidelines.	67 (17.3%)
I use alcohol-based hand rub before and after handling laboratory equipment and samples	195 (50.3%)
I properly discard biomedical waste in the right color-coded bins.	288 (74.2%)
I wash my hands after handling potentially infectious materials or biomedical waste.	319 (82.2%)
I wear personal protective equipment when collecting and working with samples.	258 (66.5%)
I organize my workplace from the clean side to the dirty side in the biosafety cabinet.	225 (58.0%)
I handle biological spills by following proper procedures.	208 (53.6%)
I change gloves whenever they are contaminated or their integrity is compromised.	250 (64.4%)
I wear an eye shield when performing activities that involve the splash or spray of hazardous materials.	174 (44.8%)
I use biosafety cabinet when handling microbiological sample and conducting relevant test.	217 (55.9%)

Table 8: Correlation coefficient and p-value for knowledge, attitude and practices scores for laboratory biosafety (n=388)

Variables	Correlation coefficient, r	*p*-value	Inference
Knowledge score- Attitude score	+0.046	0.371	Negligible and not statistically significant correlation
Knowledge score- Practice score	+0.135	**0.008**	Weak positive and statistically significant correlation
Attitude score- Practice score	+0.417	**<0.001**	Weak positive and statistically significant correlation

**Significant results (p <0.05) were bold.*

Association between knowledge level and demographic factors

The study found no significant association between the knowledge of laboratory biosafety and students' gender or faculty. However, knowledge levels varied significantly

across different years of study (p<0.001) (Table 9). Post hoc Bonferroni tests revealed significant differences between first and final year (p = 0.019), second and third year (p = 0.002) and second and final year (p<0.001) (Table 10). These findings suggest that biosafety knowledge was determined more by academic progression than by demographic factors.

Table 9: Association analysis between knowledge scores and demographic outcomes

Parameter	n	Knowledge score		Statistics (df)	*p*-value
		Mean	SD		
Gender				t-statistics	
Male	45	7.53	1.34	1.328	0.185
Female	343	7.23	1.44	(386)	
Faculty					
KAHS	243	7.34	1.41	Chi-square statistics 2.173 (5)	0.825
KOD	13	7.46	1.13		
KOM	36	7.19	1.62		
KON	44	7.16	1.26		
KOP	18	7.11	1.71		
KOS	34	7.00	1.56		
Year of study					
First year	80	7.11	1.47	F-statistics 9.275 (3, 384)	**<0.001**
Second year	129	6.84	1.53		
Third year	79	7.54	1.36		
Final year	100	7.73	1.11		
Briefing status					
Received briefing	297	7.37	1.35	t-statistics 2.48 (386)	**0.013**
Did not received briefing	91	6.95	1.61		

Table 10: Multiple comparisons (Post Hoc) of knowledge scores regarding laboratory biosafety using Bonferroni test for different year of study

Variable	Year of study	Year of study	*p*-value
Knowledge score	First year	Second year	0.977
		Third year	0.299
		Final year	**0.019**
	Second year	Third year	**0.002**
		Final year	**<0.001**
	Third year	Final year	1.000

**Significant results (p <0.05) were bold.*

DISCUSSION

Laboratory biosafety is essential for health sciences students, either undergraduate or postgraduate level, who handle biological samples, including hazardous and infectious materials, genetically modified specimens and related waste. Since students frequently work with such materials during lab sessions and research, understanding biosafety guidelines helps prevent lab accidents, protecting both individuals and the surrounding environment.

This study assessed awareness of laboratory biosafety among undergraduate students from various faculties and academic programmes related to health sciences. Results indicated that most students had a moderate level of knowledge, with an average score of 7.27 out of 10, consistent with the moderate-range framework established by Mehta & Diwakar (2021). These findings align with a similar 2021–2022 study in Vietnam, where 69% of medical laboratory students demonstrated basic lab safety awareness (Ngoc et al., 2024). The reason for this gap may be due to a lack of training on laboratory biosafety. This fact is supported by a study conducted by Keshan et al. (2020) in which they found a significant difference between pre- and post-training knowledge scores on biosafety (p<0.001). This implies that knowledge could be improved by proper training.

This study evaluated laboratory biosafety awareness among undergraduate students who frequently engaged in biological laboratory work. This assessment focused not only on their knowledge of biosafety but also their attitudes and practices related to it. These aspects were examined across five fundamental areas including hygiene, waste management, emergency equipment handling, PPE usage, and biosafety cabinet operation.

Among the knowledge-based questions, the highest score was recorded for the PPE-related item, which is the purpose of wearing eye shields. This suggests a strong understanding among students regarding the importance of this particular item. However, their understanding of PPE usage was inconsistent; while they recognised the importance of eye protection, their responses faltered on other PPE-related questions, such as identifying the correct PPE for handling pathogens or selecting the appropriate glove type.

Regarding the correct PPE items for pathogen handling, most students chose lab coats, face masks and gloves, but frequently omitted safety goggles. This oversight may be attributed from limited exposure to high-risk lab sessions involving pathogenic strains, reducing the perceived need for goggles. This assumption is supported by the low percentage (44.8%) of students who wore eye shields when handling hazardous materials, despite the fact that 80.7% acknowledge their importance.

The gap in practical compliance may reflect varying levels of lab experience across student cohorts. For instance, first-year students typically engage in low-risk lab sessions, resulting in reduced familiarity with full PPE requirements.

As a result, even though safety goggles are available, their perceived necessity remains minimal. Furthermore, since laboratory assistants typically prepare the chemicals and handle tasks involving potential splashes, students have even fewer opportunities to use safety goggles. This limited exposure to the use of eye shields and other PPE may affect students' practical habits, as supported by Rofifa et al. (2019) which suggest that PPE usage tends to increase with greater familiarity and experience in laboratory settings. Inadequate knowledge might be resulted from insufficient safety equipment in the lab, as reported by Goh (2018), where lack of safety goggles in the lab led to a decline in the awareness level due to rare use of the item in routine lab sessions.

Another PPE-related knowledge gap appeared in glove selection, where many respondents incorrectly chose 'plastic glove' over 'non-sterile glove'. This may reflect confusion between common latex gloves and plastic alternatives. Despite this, students demonstrated strong adherence to PPE principles in practice, recognising their importance when handling potentially hazardous materials. It can be evidenced by the fact that many of them change gloves when they become contaminated or when their integrity was compromised. This response demonstrated that the students strongly believed gloves could protect them from infection, although they were unaware of the type of gloves that should be used in the lab. The students' gap in theoretical knowledge likely stems from the routine practice of using provided gloves (e.g., latex) in the lab without receiving formal information about their material composition and the alternatives, leading them to use the available gloves without considering their specific type or properties. This may be due to lack of training, as proven by Mehta and Diwakar (2021), who reported that knowledge improved after training sessions. The data reveals that 23.5% of the students never receive a briefing, and only 17.1% of the students do receive it annually. This suggests that laboratory biosafety training is still inadequate, leading to insufficient knowledge. All of these responses indicated that positive practices and attitudes towards PPE commitment were interrelated but not linked to the knowledge component. This assumption was supported by the moderate positive correlation between attitude and practice, whereas knowledge showed only a weak or non-significant correlation with them.

Regarding the key aspect of waste management, nearly 95% of students knew the yellow bin is the appropriate container for discarding biohazardous waste. In addition to recognising its importance, they also demonstrated this understanding in practice, with 74.2% reporting that they consistently dispose biological waste correctly after lab work. This positive behaviour may be attributed to the availability and clear labelling of bins in the laboratory, regular briefings and reminders from lab personnel, and prior training or orientation. Similar findings have been observed in other studies, such as those involving students in Uganda (Padde et al., 2022), Saudi Arabia (Abu-Siniyeh & Al-Shehri, 2020) and South Delhi (Singh, 2022), albeit with slight variations in percentages across studies. The latter study demonstrated that students consistently practised proper waste management techniques after attending biomedical waste management class.

In contrast, knowledge related to handling biological spills was notably lower, with only 33% of students answering correctly. Over half either provided incorrect answers or were unsure. Despite this, most students believed that following proper spill management procedures could reduce the risk of infection spread. In fact, over half claimed that they always follow proper procedures in such situations. This discrepancy raises concerns about whether students may be unknowingly practising incorrect steps while believing they are following the right procedures. A closer look at the responses revealed that many students believed the correct approach was to directly apply disinfectant before wiping and discarding it, often omitting the crucial step of first covering the spill before applying disinfectant. This suggests gap training or hands-on experience, possibly because the spill incidents are rare or typically handled by lab managers. Proper spill management is critical for minimising lab-acquired infections (LAIs), one of the most significant risks in laboratory work. These infections could escalate to a global scale if improper cleaning of biohazardous materials becomes a consistent practice (Demaneuf, 2020). Notably, students appeared to be aware of the potential for infection from contaminated biological materials and acknowledged the risk of greater spread in crowded environments. Crowded workspaces can increase the risk of transmission infection due to close contact and the chances of encountering items strewn across the table (Alshammari et al., 2021).

Encouragingly, the present study revealed that the students possessed relatively high levels of knowledge and awareness regarding the procedures and importance of proper handwashing, whether using soap or alcohol-based hand sanitisers. While most students reported that they frequently washed their hands after lab sessions, many did not always use soap. Furthermore, the use of alcohol-based hand rubs after lab sessions was less common among the students in this study, contrasting with findings from India, where students demonstrated higher adherence to such practices (Modi et al., 2017). Inconsistent training may contribute to poor compliance,

as evidenced by Dutta et al. (2020), who found a strong link between hand hygiene training and practice. Students who received training were more likely to wash their hands consistently after removing gloves and completing lab sessions. The practices of hand hygiene are crucial in preventing incidents of cross-contamination caused by unclean hands during lab activities, regardless of the level of awareness and belief. Therefore, the biosafety unit responsible for these protocols should reinforce the importance of hand hygiene through regular training and highlight the potential consequences of improper practices. This proactive approach not only prevents harm to individuals but also helps minimise risks and prevent broader impact at the community level.

Another critical aspect of biosafety regulation in the biology laboratories is adherence to proper biosafety cabinet (BSC) usage procedures. A high percentage of respondents demonstrated good understanding of the function of a BSC, which is to protect personnel, environment and products when handling risk items, and they believed in the importance of the biosafety cabinet in ensuring safety. Practically, half of the students claimed that they use biosafety cabinet when handling microbiological samples, even they manage to organise the workplace appropriately, which is to separate the clean and dirty sides in the biosafety cabinet to reduce cross contamination during the lab activities. This score is among the lowest scores obtained for the questions in the survey. This low score could be due to the nature of experiments set up for undergraduate students who frequently practise on open bench due to low-risk research activities (Meechan & Paul, 2020). However, the importance of using biosafety cabinet should get appropriate concern by the institution, as previous case have shown a great health impact on the lab users who are exposed to pathogenic Brucellosis's infectious agents due to failure to adhere to the utilisation of biosafety cabinet (Sam et al., 2012). In essence, students' knowledge and attitude regarding BSC were good; however, their practice did not align with these strengths.

The final key area of biosafety assessed was the use of fire extinguishers as emergency equipment in the lab. Approximately 66% of students correctly identified the Class ABC fire extinguisher as the appropriate tool for handling fire incidents in the lab. However, a significant number of students, which is 120 of them, either provided incorrect answers or were unsure. Despite this gap, nearly 70% strongly agreed that everyone should know how to use a fire extinguisher. Yet, almost half reported that they never or rarely used one in the laboratory. This situation is similar to findings from Saudi Arabia, where students lacked opportunities to use fire extinguisher simply because no real fire incidents requiring them to use them (Abu-Siniyeh & Al-Shehri, 2020). In emergency situations such as fires, students' ability to respond effectively using a fire extinguisher is very crucial. To bridge this gap, institutions should implement regular fire drills and hands-on training, allowing students to learn, observe and practices proper fire extinguisher use.

The study also examined potential factors influencing students' understanding of laboratory biosafety. Key variables assessed included gender, academic programme, year of study, and prior exposure to laboratory biosafety training. The results revealed significant differences in knowledge scores based on academic year and pre-laboratory training.

Multiple comparisons analyses showed significant score disparities between first-year and final-year students, as well as between second-year students and their senior counterparts (third-year and above). This finding aligns with the initial assumption that junior students possess less laboratory experience than seniors. Furthermore, most academic programmes typically schedule off-campus practical sessions for third-year students and above, suggesting that senior students may acquire both formal and informal biosafety knowledge through their academic progression. This correlation has been consistently reported in previous studies. For instance, research demonstrated that while first-year students often struggle to identify chemical and biohazard symbols, internship students perform significantly better (Walters et al., 2017; Abu-Siniyeh & Al-Shehri, 2020; Geraldez et al., 2023; Ngoc et al., 2024).

Prior training emerged as a primary determinant of biosafety comprehension. The current findings corroborate earlier studies emphasising the importance of early biosafety instruction (Abu-Siniyeh & Al-Shehri, 2020; Dutta et al., 2020; Mehta & Diwakar, 2021; AlWaqfi et al., 2022). While traditional methods such as passive instruction and verbal briefing demonstrate moderate effectiveness, active students' engagement during safety training yields significantly better knowledge retention. Recent innovative approaches show particular promise, as demonstrated by Li et al. (2024), who reported medical students exhibited enhanced biosafety understanding through a blended experiential and mobile learning approach using WeChat. Given students' increasing technology proficiency, incorporating AI and similar technologies could potentially optimise biosafety compliance and adaptation in laboratory settings.

CONCLUSION

This study evaluated laboratory biosafety awareness among undergraduate students at IIUM who regularly engage in biological laboratory work. Among 388 students, the majority demonstrated moderate biosafety knowledge levels. While students showed satisfactory understanding of five key elements, particularly PPE utilisation, waste management and hygiene practices, this knowledge inconsistently translated into practical application. The concerning disconnect between theoretical knowledge and actual practice highlights a critical area for intervention. These findings underscore the urgent need to bridge this implementation gap, particularly for students who spend significant time in laboratory settings. Future biosafety initiatives should prioritise behavioural compliance alongside knowledge acquisition, ensuring that safety protocols become ingrained in routine laboratory practice. Furthermore, the study reveals that awareness alone is insufficient for ensuring biosafety compliance. Institutional efforts should focus on developing targeted training programmes that emphasize hands-on application, regular reinforcement of safety protocols and accountability measures to transform knowledge into habitual safe practices. This approach would be particularly impactful in academic settings where students develop their foundational laboratory habits that may persist throughout their professional careers.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Abu-Siniyeh, A., & Al-Shehri, S. S. (2020). Safety in Medical Laboratories: Perception and Practice of University Students and Laboratory Workers. *Applied Biosafety, 26*(S1). https://doi.org/10.1089/apb.20.0050

Ahmed, H. K., Eisa, I. M., Abdallah, E. I., Hamouda, D. G., Omer, A. E., & Eltayeb, L. B. (2022). Appraisal of Biosafety Measures in Governmental Medical Laboratory Personnel: Knowledge, Attitude, Practice (KAP) Study. *Journal of Biochemical Technology, 13*(3), 13-18. https://doi.org/10.51847/EOHG7NGYvX

AlShammari, W., Alhussain, H., & Rizk, N. (2021). Risk Management Assessments and Recommendations Among Students, Staffs, and Health Care Workers in Educational Biomedical Laboratories. *Risk Management and Healthcare Policy, Volume 14*, 185–198. https://doi.org/10.2147/rmhp.s278162

AlWaqfi, A. S. A., Guan, N. Y., Ying, L. P., & Tamrin, S. B. M. (2022). Factors Associated with Knowledge, Attitude and Practices on Fire Safety and Its Prevention Among Hostel Occupants in a Higher Learning Institution. *Malaysian Journal of Medicine & Health Sciences*, 18.

Bayot, M. L., & Limaiem, F. (2019, January 31). *Biosafety Guidelines*. Nih.gov; StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK537210/

Burdass, D., Grainger, J., Hurst, J., & Society For General Microbiology. (2005). *Basic Practical Microbiology : a Manual*. Society For General Microbiology. https://microbiologysociety.org/static/uploaded/23cbf9c5-f8c8-4f91-b092a4ad819e6357.pdf

Chen, X., Lin, G., Wang, K., Yang, T., Feng, Y., Wang, H., Li, D., & Fu, G. (2023). Current challenges of university laboratory: Characteristics of human factors and safety management system deficiencies based on accident statistics. *Journal of Safety Research, 86*, 318–335. https://doi.org/10.1016/j.jsr.2023.07.010

Demaneuf, G. (2020, November 27). *A review of SARS Lab Escapes in 2003-2004*. ResearchGate. https://doi.org/10.5281/zenodo.4293257

Dutta, G., Singh, Tg., & Kumar, T. (2020). Knowledge and practice of hand hygiene among undergraduate students and junior doctors in the Regional Institute of Medical Sciences, Imphal. *Journal of Family Medicine and Primary Care*, 9(9), 4741. https://doi.org/10.4103/jfmpc.jfmpc_794_20

El-Gilany, A.-H. ., El-shaer, S., Khashaba, E., El-dakroory, S. A., & Omar, N. (2017). Knowledge, attitude, and practice (KAP) of "teaching laboratory" technicians towards laboratory safety and waste management: a pilot interventional study. *Journal of Hospital Infection, 96*(2), 192–194. https://doi.org/10.1016/j.jhin.2017.02.007

Geraldez, Florence, K., Aguilar, C., Toledo, I. W., & Deocades, E. I. (n.d.). Knowledge, Attitude, and Practices Towards Laboratory Safety Measures Among Notre Dame of Marbel University Medical Technology Students. *Psych Educ, 2023*(8), 2023–2706. https://doi.org/10.5281/zenodo.7928184

Goh, C. T. (2018). Assessing Awareness on Laboratory Safety: A Case Study in Pahang, Malaysia (Penilaian Kesedaran Keselamatan Makmal: Kajian Kes di Pahang,

Malaysia). *Jurnal Pendidikan Malaysia*, 43(02). https://doi.org/10.17576/jpen-2018-43.02-07

Gugnani, H. C., & Randhawa, H. S. (2020). Laboratory-acquired fungal infections: a review. *Archives of Microbiology & Immunology, 4*(2), 51-56. https://doi.org/10.26502/ami.93650044

Keshan, P., Rastogi, A., Aggarwal, S., Nigam, A., Kapila, R., & Syed, S. (2020). Effect of one-day training on Knowledge related to Biosafety and waste management among life-science Students. *Indian Journal of Community Health*, *32*(4), 694–698. https://doi.org/10.47203/IJCH.2020.v32i04.014

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement, 30*(3), 607–610. https://doi.org/10.1177/001316447003000308

Li, Q., Zhao, J., Yan, R., Gao, Q., Zhen, Y., Li, X., Liang, Y., Min, S., & Yang, L. (2024). WeChat mini program in laboratory biosafety education among medical students at Guangzhou Medical University: a mixed method study of feasibility and usability. *BMC Medical Education*, *24*(1). https://doi.org/10.1186/s12909-024-05131-9

Lim, P. L., Kurup, A., Gopalakrishna, G., Chan, K. P., Wong, C. W., Ng, L. C., Se-Thoe, S. Y., Oon, L., Bai, X., Stanton, L. W., Ruan, Y., Miller, L. D., Vega, V. B., James, L., Ooi, P. L., Kai, C. S., Olsen, S. J., Ang, B., & Leo, Y.-S. (2004). Laboratory-Acquired Severe Acute Respiratory Syndrome. *New England Journal of Medicine, 350*(17), 1740–1745. https://doi.org/10.1056/nejmoa032565

"Meechan, P. J., & Potts, J. (2020). *Biosafety in microbiological and biomedical laboratories*. https://www.cdc.gov/labs/pdf/SF__19_308133-A_BMBL6_00-BOOK-WEB-final-3.pdf

Mehta, A., & Diwakar, M. Kr. (2021). Impact of Educational Intervention on Knowledge, Attitudes, and Practices of Laboratory Technicians regarding Universal Work Precautions at a Tertiary Health Care Center in Central India. *Applied Biosafety, 26*(S1), S-27-S-33. https://doi.org/10.1089/apb.21.902607

Modi, P. D., Kumar, P., Solanki, R., Modi, J., Chandramani, S., & Gill, N. (2017). Hand Hygiene Practices Among Indian Medical Undergraduates: A Questionnaire-Based Survey. *Cureus, 9*(7). https://doi.org/10.7759/cureus.1463

Ngoc, T., Trung, N., Xuan, T., & Thu, T. (2024). Biosafety knowledge and perception among medical laboratory students: a cross-sectional study at a medical university in Vietnam. *International Journal of Occupational Safety and Health, 14*(1), 7–16. https://doi.org/10.3126/ijosh.v14i1.56395

Padde, J. R., Akiteng, W., Edema, W., Atiku, S. M., Tibyangye, J., Tekakwo, J., Andruga, C., Hope, D., Musinguzi, B., Gesa, J. B., Amadile, L., & Agondua, R. (2022). Assessment of Biosafety and Biorisk Management Practices among Medical Laboratory Students in Two Institutions in Uganda. *Biosafety and Health, 4*(06), 399–405. https://doi.org/10.1016/j.bsheal.2022.08.005

Peng, H., Bilal, M., & Iqbal, H. (2018). Improved Biosafety and Biosecurity Measures and/or Strategies to Tackle Laboratory-Acquired Infections and Related Risks. *International Journal of Environmental Research and Public Health, 15*(12), 2697. https://doi.org/10.3390/ijerph15122697

Rofifa, A. T., Alayyannur, A., & Haqi, D. N. (2019). Analysis of Factors Related to Use of Personal Protective Equipment (PPE) in Laboratory. *Malaysian Journal of Medicine and Health Sciences, 15*(3), 2636–9346

Sam, I. C., Karunakaran, R., Kamarulzaman, A., Ponnampalavanar, S., Omar, S. S., Ng, K. P., ... & AbuBakar, S. (2012). A large exposure to Brucella melitensis in a diagnostic laboratory. *Journal of Hospital Infection, 80*(4), 321-325. https://doi.org/10.1016/j.jhin.2011.12.004

Schober, P., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia and Analgesia, 126*(5). https://doi.org/10.1213/ANE.0000000000002864

Singh, A. K. (2022). A Comparative Study on Perception about Biomedical Waste Management among the Undergraduate and Postgraduate Students in a Tertiary Care Teaching Hospital in South Delhi. *The Journal of Communicable Diseases, 54*(4), 1–6. https://doi.org/10.24321/0019.5138.202295

Walters, A. U. C., Lawrence, W., & Jalsa, N. K. (2017). Chemical laboratory safety awareness, attitudes and practices of tertiary students. *Safety Science, 96*, 161–171. https://doi.org/10.1016/j.ssci.2017.03.017

World Health Organization. (2010). *WHO best practices for injections and related procedures toolkit* (No. WHO/EHT/10.02). World Health Organization.

World Health Organization. (n.d.). *WHO EMRO | Biosafety | Health topics*. World Health Organization - Regional Office for the Eastern Mediterranean. Retrieved November 5, 2023, from https://www.emro.who.int/health-topics/biosafety/index.html

REVIEW ARTICLE

Maternal Exposure to Polycyclic Aromatic Hydrocarbon (PAH) and Its Effects on Pregnancy Outcomes: A Systematic Review

Nabilah Husna Mohd Nazumi[1], Norhidayah Ahmad[1,*]

[1]Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia.

Keywords:
polycyclic aromatic hydrocarbon; pregnancy outcome; intrauterine growth; maternal exposure; air pollution

ABSTRACT

Background: In the natural environment, polycyclic aromatic hydrocarbons (PAHs) are predominantly associated with fine particulate matter ($PM_{2.5}$) suspended in the ambient air. These tiny particles can be inhaled by humans and subsequently deposited in the respiratory system, where they can penetrate the bloodstream and accumulate in target organs, including the placenta. This condition resulted in placental dysfunction, which can lead to adverse pregnancy outcomes. Therefore, this study aims to systematically review the scientific published articles on maternal exposure towards PAHs during pregnancy and the effect on adverse pregnancy outcomes. **Methods:** This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A systematic article search was performed using three online databases: PubMed, Scopus, and ProQuest to identify articles published from 2000 to 2022. A thorough article search identified eight studies suitable for inclusion in this review, as well as two additional studies discovered through the snowballing technique. All the included studies underwent quality assessment by using the NIH quality assessment tool. **Results:** A total of 948 articles were screened, of which ten full-text articles met the inclusion criteria for final evaluation. The most frequently reported adverse pregnancy outcomes across the selected studies were fetal growth restriction, including reduced birth weight and birth length, decreased head circumference, small for gestational age (SGA) and preterm delivery. PAH-DNA adducts in the cord blood were significantly associated with prenatal exposure to PAHs. PAHs exposure appears to have the most adverse impact on embryonic development during the first trimester. Seven articles recorded high quality (score range from 11 to 12.5 points) and only three articles recorded fair quality (score range from 8.5 to 10.5 points). **Conclusion:** Our findings indicate that maternal exposure to PAHs could negatively impact pregnancy outcomes. While this review enhances the understanding of the observed relationship, only a limited number of studies have explored the underlying biological mechanism. The appropriate measures to protect and limit the maternal exposure to PAHs, particularly in the first trimester, must be implemented. These include providing clinical advice to the expectant mothers on reducing exposure to tobacco smoke and poor air quality, encouraging appropriate dietary modifications, as well as reinforcing governmental efforts to strengthen environmental regulations.

INTRODUCTION

Polycyclic aromatic hydrocarbons (PAHs), also known as polynuclear aromatic hydrocarbons (PNAs), refer to a class of chemical compounds that consist only of hydrogen and carbon (Abdel-Shafy & Mansour, 2016). Due to their tiny size, PAHs that attach to $PM_{2.5}$ can penetrate the bloodstream and accumulate in the target organ, including the placenta, which can cause placental dysfunction and lead to adverse pregnancy outcomes (Abdel-Shafy & Mansour, 2016; Perera et al., 2005a). Several PAHs have been shown to cause cancer, including naphthalene, chrysene, B[*a*]P, dibenzo[*a,e*]pyrene, dibenzo[*a,l*]pyrene and anthanthrene (Dybing et al., 2013).

PAHs can enter the human body through the lungs during breathing, by drinking contaminated water, and by consuming food, dirt, or dust particles that contain PAHs. PAHs can reach the human body through multiple routes of exposure, under normal environmental conditions, when the body comes into contact with dirt or dust that contains PAH (Mohd Radzi et al., 2016).

The elevated levels of B[a]P exposure during pregnancy lead to birth abnormalities and low birth weight, as shown in mice's laboratory experiments (Kim et al., 2013). Studies by the U.S. Center for Children's Environmental Health have found that maternal exposure to PAHs was correlated with adverse pregnancy outcomes such as reduced birth weight, premature birth and impaired infant growth (Perera et al., 2005b).

* *Corresponding author.*
E-mail address: hidayahahmad@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

Elevated levels of B[a]P were associated with the decline in the thickness of the placenta and were able to impair the placental structure and function, and lead to reproductive adverse consequences by crossing the placental barrier (Al-Saleh et al., 2013). Elevated levels of PAH-DNA adducts are related to the restriction of fetal and child growth (Tang et al., 2006). The accumulation of PAHs more easily occurred in fetal tissue than in the placenta and maternal tissue. The difference in enzyme function resulted in lower metabolism of PAHs in the fetal tissue (Anderson et al., 2000).

Fetal vulnerability to high exposure to PAHs varies at different gestational ages throughout the gestational period (Wang et al., 2014). Other epidemiologic studies have found that the rate of fetal growth is regulated throughout the first trimester, resulting in a progressively greater deficit as the pregnancy progresses (Milani et al., 2005; Smith, 2004; Neufeld et al., 1999). Implantation and initial placentation during the first trimester are two critical events that are vulnerable to placental disruption (Erickson & Arbour, 2014). PAH exposure in the early stage of gestation can disrupt the early proliferation of trophoblast endovascular cells and prevent them from infiltrating into the fetal envelope. This condition leads to further disturbance of placental development and may restrict the supply of oxygen and nutrients to the fetus (Smith, 2004; Chaddha et al.,2004; Ritz & Wilhelm, 2008).

Compared to adults, developing fetuses and also neonates are more susceptible to the effects of the toxicology of PAHs, and when compared to non-pregnant women, pregnant women are at a higher risk (Jedrychowski et al., 2013). They are especially at risk because, from mother to fetus, PAHs can pass through the placenta (Polanska et al., 2014). Given the susceptibility of developing fetuses and neonates, this systematic review aims to examine maternal exposure to PAHs during pregnancy, its association with adverse pregnancy outcomes, and to identify the gestational period of greatest vulnerability to PAHs exposure.

MATERIALS AND METHODS

Study Design

This study was carried out based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline.

Search Strategy

A bibliographic search was conducted using the online databases of PubMed, Scopus, and ProQuest, accessed through the IIUM Library account, and employing a backwards snowballing technique. Keywords that were used include “maternal exposure”, “prenatal exposure”, “polycyclic aromatic hydrocarbon*”, “adverse pregnancy outcome*”, “intrauterine growth”, “fetal growth retardation”, “pregnancy”, “cord blood”, “maternal serum”, “placenta”, “biological samples”. The Boolean operators of the connective terms, such as AND, OR, and NOT, have been used in the search strategy to obtain comprehensive results.

Selection Criteria

From the identified articles, studies that met the following eligibility criteria were selected based on: (i) articles published in the English language with full-text access from year 2000 or later; (ii) human epidemiological studies with several study designs (cohort, cross-sectional and case control studies) and (iii) studies reported on maternal exposure to PAHs during pregnancy and the effect on pregnancy outcomes. Retrieved articles that did not meet these criteria were excluded from the list. Only the final shortlisted articles went through the data analysis process and quality assessment procedures.

Article Screening and Data Extraction

Two reviewers independently performed article selection to ensure relevance and adherence to the criteria of the study. Any discrepancies between reviewers were resolved through discussion until consensus was achieved. The selection of articles based on the PRISMA guideline comprises four phases: (i) identification, (ii) screening, (iii) eligibility and (iv) included studies.

For the first level of screening, the titles and abstracts of the articles were reviewed, and duplicates of research articles across databases were removed. The articles that passed the initial screening were further evaluated based on the inclusion and exclusion criteria. Then, the data from the selected articles, including study objectives, methodology, number of participants, results, discussion of the findings, mechanism of action, and primary outcomes of the studies, were extracted and reviewed (as shown in Table 1).

Quality Assessment

The quality assessment of the selected studies was performed using the National Institutes of Health (NIH) quality assessment tool, designed for evaluating cohort, cross-sectional and case-control study designs (NIH, 2014). This tool assesses key aspects of study quality, including the strength of the study objectives, the credibility of the

study, the findings and discussion, and the participant's evaluation or the population involved in the study. The quality assessment comprises 14 questions for cohort and cross-sectional studies, and 12 questions for case-control studies. Each question can be scored as 1 for "yes", 0.5 for "no" and 0 for "not applicable", "cannot be determined" or "not reported". The total scores were summed up and categorised as follows: good quality (11–14 points), fair quality (7.5–10.5 points), and poor quality (0–7 points).

RESULTS

Bibliographic Search

The articles search retrieved 948 articles from the databases, and 12 additional articles were retrieved using the snowballing method. After removing the duplication, 607 articles were screened based on titles and abstracts. Then, 70 articles were assessed by referring to their full texts to examine the eligibility of the criteria based on the inclusion and exclusion criteria. Sixty articles that did not fulfil the inclusion criteria were excluded from the final selection. Finally, 10 articles were included in this systematic review. The final number of studies identified, included, and eliminated was summarised and recorded in the PRISMA Flow Diagram in Figure 1.

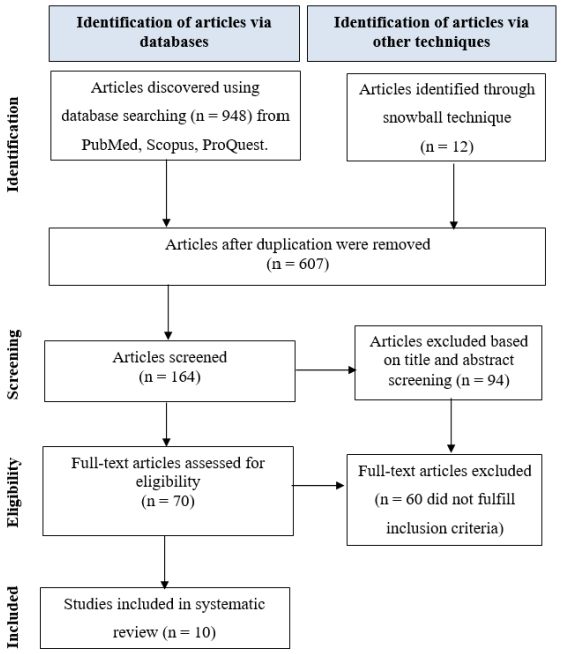


Figure 1: PRISMA Flow diagram for article selection

Overview of the Selected Articles

Table 1 presents a summary of the characteristics and principal findings of the selected articles in this systematic review. The selected studies have been published between 2000 and 2018. Among the 10 selected articles, eight were prospective cohort studies and two were case-control studies. The studies by Agarwal et al. (2018), Jedrychowski et al. (2017), Polanska et al. (2010), Choi et al. (2008), Singh et al. (2008), and Perera et al. (2005) have addressed maternal exposure to PAHs and adverse pregnancy outcomes, whereas studies by Jedrychowski et al. (2013), Zhang et al. (2016), and Gladen et al. (2000) primarily focused on the accumulation of PAHs in biological samples. Only a study by Choi et al. (2012) specifically examined the most vulnerable period to PAHs exposure during pregnancy.

Accumulated PAHs in Biological Samples

The studies in Poland utilised data from a previously established birth cohort of infants in Krakow (Jedrychowski et al., 2017; Jedrychowski et al., 2013; Choi et al., 2012) Four studies analysed and measured PAHs level through the sample that have been collected from the personal air monitor (Jedrychowski et al., 2017; Jedrychowski et al., 2013; Choi et al., 2012; Choi et al., 2008), three studies assessed the level of PAHs through placental tissue samples (Agarwal et al., 2018; Singh et al., 2008; Gladen et al., 2000), one study observed PAHs level within variety biological samples (maternal serum, umbilical cord serum and placenta) (Zhang et al., 2016), one study through umbilical cord blood and maternal blood (Perera et al., 2005) and one study through concentration of 1-hydroxypyrene (1-HP) in urine (Polanska et al., 2010).

All seven PAHs were detected in all placentas, except benzo[a]pyrene, which was undetected in one placental sample (Gladen et al., 2000). Maternal serums had the highest concentration of total PAHs levels (1290 ng g-1 lipid), followed by umbilical cord serums (1150 ng g-1 lipid) and lowest in placenta samples (673 ng g-1 lipid) (Zhang et al., 2017). B[a]P exposure during pregnancy causes DNA damage in the biological samples of maternal and cord blood (Jedrychowski et al., 2013). PAH-DNA adducts in the cord blood were shown to be significantly related to tertiles of B[a]P exposure (non-parametric trend z = 3.50, p <0.001) (Jedrychowski et al., 2013).

Maternal Exposure to PAH and Effects on Pregnancy Outcomes

A case-control study conducted in India by Agarwal et al. (2017) found that preterm birth groups (0.485 ± 0.675 µg/L) had significantly higher levels of B[a]P in the placental samples, compared to full-term delivery groups (0.124 ± 0.436 µg/L). Similarly, a case-control study by Singh et al. (2008) reported that mothers with preterm deliveries had significantly higher carcinogenic PAH benzo(b)fluoranthene, with concentrations of 61.91 ± 12.43 ppb in the preterm birth group compared to 23.84 ± 7.01 ppb in the full-term delivery group.

Table 1: Study characteristics and principal findings of selected studies

Author	Study Characteristics	Methodology	Principal Findings
Agarwal et al. (2018)	Case-control study. Sample population: 55 controls (full-term delivery), 29 cases (preterm delivery). Not trimester-specific exposure. Measure the relationship between PAH exposure and preterm delivery.	Interview, 30g of placental tissues, hospital records.	**Level of B[*a*]P** Preterm birth groups: (0.485 ± 0.675 µg/L) Full-term delivery groups: (0.124 ± 0.436 µg/L). **Level of GSH in cases than in controls** Significantly lower ($p < 0.05$) **Level of MDA in cases than in controls** Significantly higher ($p < 0.05$)
Zhang et al. (2016)	Prospective cohort study. Sample populations: 95 participants. Not trimester-specific exposure. Measure PAH in biological samples.	Questionnaire, 286 individual samples of maternal blood, umbilical cord blood and placentas.	**ΣPAH15** Maternal serum: 1290 ng g^{-1} lipid Cord serums: 1150 ng g^{-1} lipid Placentas: 673 ng g^{-1} lipid **B[a]Peq** Maternal serum: 22.4 ng g^{-1} lipid Cord serums: 20.4 ng g^{-1} lipid Placenta: 3.2 ng g^{-1} lipid
Jedrychowski et al. (2017)	Prospective cohort study. Sample populations: 455 participants. 2nd trimester. PAH exposure and birth weight, birth length and head circumference.	Questionnaire, wearing a personal air monitor during the 2nd trimester for 2 days. Maternal and child hospital records.	**Birth weight (g)** PAH: (β = -0.20, $p = 0.004$) $PM_{2.5}$: (β = -0.02, $p = 0.757$) **Birth length (cm)** PAH: (β = -0.17, p = 0.025) $PM_{2.5}$: (β = -0.04, p = 0.453) **Head circumference (cm)** PAH: (β = -0.13, p = 0.086) $PM_{2.5}$: (β = -0.06, p = 0.277)
Choi et al. (2012)	Prospective cohort study. Sample population: 344 participants. 2nd trimester ($n = 344$), each trimester ($n = 77$). PAH exposure and reduction in FGR, birth weight, birth length and birth head circumference.	Questionnaire and wearing a personal air monitor during the 2nd trimester and each trimester for 2 days. Medical record and calculation of fetal growth ratio (FGR, %) of newborns.	**Reduction in each trimester** **FGR** 1st: (-3%, 95% CI, -5 to -0%) 2nd: (-1%, 95% CI, -2% to 1%) 3rd: (0%, 95% CI, -3% to 1%) **Birthweight** 1st:(-105 g, 95% CI, -188 to -22g) 2nd: (-36 g, 95% CI, -76 to 4 g) 3rd: (-44 g, 95% CI, -123 to 36 g) **Birth length** 1: (-0.78 cm, 95% CI, -1.30 to -0.26 cm) 2nd : (-0.24 cm, 95% CI, -0.49 to 0.01 cm) 3rd : (-0.57 cm, 95% CI, -1.07 to -0.07 cm) **Birth head circumference** 1st: (-0.11cm, 95% CI, -0.39 to 0.17 cm) 2nd : (-0.13 cm, 95% CI, -0.26 to 0.00 cm)
Polanska et al. (2010)	Prospective cohort study. Sample population: 449 mother-child pairs. 2nd trimester. PAH exposure and birth weight, birth length, head circumference, ponderal index and cephalisation index.	Questionnaire, urine samples of 20-24 weeks pregnant women, saliva samples once each trimester. medical records of newborns.	**Birth weight (g)** (β = -158.3, p = 0.01) **Birth length (cm)** (β = -0.70, p = 0.05) **Head circumference (cm)** (β = -0.50, p = 0.05) **Ponderal index (g/cm^3)** (β = -0.02, p = 0.60) **Cephalisation index (cm/g)** (β = 4.2, p = 0.01)

Choi et al. (2008)	Prospective cohort study. Sample population: African American (n=224) and Dominican women (n=392). 3rd trimester. PAH exposure and SGA, FGR < 85% and preterm delivery.	Questionnaire, wearing a personal air monitor for 2 days. Medical record.	**Preterm delivery** African American (n = 205): (OR = 4.676, *p* = 0.001) Dominican (n = 351): (OR = 0.523, *p* = 0.229) **SGA** African American (n = 201): (OR = 1.94, *p* = 0.03) Dominican (n = 336): (OR = 0.82, *p* = 0.53) **FGR** (full term only) African American (n = 194): (OR = 2.15, *p* = 0.03) Dominican (n = 341): (OR = 0.80, *p* = 0.40)
Jedrychowski et al. (2013)	Prospective cohort study. Sample population: 362 2nd trimester. PAH-DNA adducts in cord blood.	Questionnaire, wearing a personal air monitor for 2 days. Maternal blood and umbilical cord blood were collected.	**Cord blood concentrations of PAH–DNA adducts (per 10^8 nucleotides)** ≤ 1.15 ng/m^3 : (0.24, 95% CI 0.22 to 0.26) 1.16-5.69 ng/m^3 : (0.25, 95% CI 0.25 to 0.27) > 5.69 ng/m^3 : (0.30, 95% CI 0.27 to 0.32) (*z* = 3.50, *p* < 0.001) **Maternal blood concentrations of PAH-DNA adduct (per 10^8 nucleotides)** ≤ 1.15 ng/m^3 : (0.25, 95% CI 0.23 to 0.27) 1.16-5.69 ng/m^3 : (0.26, 95% CI 0.24 to 0.28) > 5.69 ng/m^3 : (0.27, 95% CI 0.25 to 0.30) (*z* = 1.63, *p* = 0.103)
Singh et al. (2008)	Case-control study. Sample population: Control (*n* = 31), cases (*n* = 29) Not trimester-specific exposure. Compare the placental PAH levels in the preterm delivery group with the control group of normal delivery.	Interview, collection of placental tissues and medical record.	**PAHs (ppb)** **Naphthalene** Full-term: 249.62 ± 47.55 Preterm: 374.32 ± 75.90 **Acenaphthylene** Full-term: 58.72 ± 42.50 Preterm: 99.09 ± 46.72 **Phenanthrene** Full-term: 377.78 ± 79.54 Preterm: 468.64 ± 105.40 **Anthracene** Full-term: 25.81 ± 8.08 Preterm: 33.26 ± 11.15 **Fluoranthene** Full-term: 208.6 ± 21.93 Preterm: 325.91 ± 45.14* **Pyrene** Full-term: 296.43 ± 91.55 Preterm: 187.24 ± 84.94 **Benzo(k)fluoranthene** Full-term: 29.85 ± 22.34 Preterm: 11.25 ± 7.15 **Benzo(b)fluoranthene** Full-term: 23.84 ± 7.01 Preterm: 61.91 ± 12.43* **Benzo(a)pyrene** Full-term: 8.83 ± 5.84

			Preterm: 13.85 ± 7.06 **Dibenzo(a,h)anthracene** Full-term: 22.03 ± 17.06 Preterm: 15.33 ± 7.06
Perera et al. (2005)	Prospective cohort study. Sample population: 369 Not trimester-specific exposure. Birth weight, birth length and head circumference.	Healthy singleton pregnant women, proximity to the WTC site, 18-39 years old, non-smokers, pregnant on 11th September 2001 and due to give birth within 41 weeks of that day.	**Natural log transformed** **Adducts in cord blood** Birth weight: (β = 0.03, *p* = 0.18) Birth length: (β = 0.001, *p* =0.96) Head circumference: (β = 0.01, *p* = 0.27) **Cord adducts x ETS** Birth weight: (β = -0.11, *p* = 0.03) Birth length: (β = 0.01, *p* = 0.53) Head circumference: (β = -0.04, *p* = 0.04)
Gladen et al. (2000)	Prospective cohort study. Sample population:200 Not trimester specific exposure PAH level in placenta and birth weight.	Questionnaire, placental samples and medical records for pregnancy and delivery.	**Median concentration (ng/g dry weight)** Benz[*a*]anthracene: 1.15 Benzo[*b*] fluoranthene: 0.86 Benzo[*g,h,i*] perylene: 0.73 Benzo[*a*]pyrene: 0.91 Chrysene: 1.38 Dibenz[*a,h*]anthracene: 0.73 Indeno[1,2,3-*c,d*]pyrene: 0.74 Total of seven PAHs: 7.36

Vulnerable Gestational Period of PAHs Exposure

A study conducted by Choi et al. (2012) was one of the earlier studies that investigated the specific period of fetal susceptibility to PAHs exposure on adverse pregnancy outcomes. This study measured exposure to eight airborne PAHs in all participants during the second trimester, and measurements were also conducted once per trimester for a subgroup. During the first trimester, one natural-log unit increase in prenatal exposure of PAHs, associated with the most significant decrease in the fetal growth ratio (FGR) (-3%, 95% CI, -5 to -0%), decrease in birth weight (-105g, 95% CI, -188 to -22g) and decrease in birth length (-0.78cm, 95% CI, -1.30 to –0.26cm) (Choi et al., 2012).

A study by Jedrychowski et al. (2017) and Choi et al. (2008) monitored PAH exposure using personal air monitoring throughout the second and third trimesters of pregnancy, respectively. Prenatal exposure to PAHs during the second trimester significantly affected fetal development, particularly birth weight. The standardised effect of PAHs on birth weight was ten times greater (β = −0.20, *p* = 0.004) compared to the effect of $PM_{2.5}$ (β = −0.02, *p* = 0.757) (Jedrychowski et al., 2017). Prenatal PAHs exposure is significantly related to the development of SGA (95% CI, 1.1-3.5; *p* = 0.03) and preterm birth (95% CI, 1.8–11.9; *p* = 0.001) among African Americans (Choi et al., 2008).

Quality Assessment of the Selected Articles

The National Institute of Health (NIH) tool has been used to assess the quality of the selected studies. The majority (seven out of eight) of prospective cohort studies (Jedrychowski et al., 2017; Jedrychowski et al., 2013; Choi et al., 2012; Polanska et al., 2010; Choi et al., 2008; Perera et al., 2005; Gladen et al., 2000), as shown in Table 2, were recorded as “good quality” with the scores ranging between 11 and 12.5. Two case-control studies (Agarwal et al., 2018; Singh et al., 2008), as shown in Table 3, were assigned a “fair quality” rating, with identical scores of 8.5. The highest score (12.5 out of 14) was obtained by Choi et al. (2012), who measured comprehensive exposure to PAH throughout the pregnancy period, with measurements taken more than once for a subset of participants for each specific trimester.

DISCUSSIONS

Maternal Exposure to PAHs and Adverse Pregnancy Outcomes

Reduction in Birth Weight, Birth Length and Head Circumference

In 2005, Perera et al. reported no significant relationship between maternal exposure to PAHs during pregnancy and fetal growth reduction (Perera et al., 2005a). This finding contrasts with their earlier study in 1998, which demonstrated a significant association between PAH-DNA adducts and birth weight (Perera et al., 1998). However, in the 2005 study, by a similar research group, a significant association was only observed when environmental tobacco smoke (ETS) exposure was included in the analysis, corresponding to a two-fold increase in the concentration of cord blood adducts (Perera et al., 2005b).

Table 2: Checklist for Quality Assessment for Prospective Cohort Study (Source: NIH Study Quality Assessment Tools)

Criteria	Prospective Cohort Study							
	Gladden et al. (2000)	**Perera et al. (2005)**	**Choi et al. (2008)**	**Polanska et al. (2010)**	**Choi et al. (2012)**	**Jedrychowski et al. (2013)**	**Zhang et al. (2016)**	**Jedrychowski et al. (2018)**
Q1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Q2. Was the study population clearly specified and defined?	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Q3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes	Yes	Yes	Yes	C/D	Yes
Q4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Q5. Was a sample size justification, power description, or variance and effect estimates provided?	No	No	No	No	No	No	No	No
Q6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	C/D	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Q7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Q8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Q9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Q10. Was the exposure(s) assessed more than once over time?	No	No	No	No	Yes	No	No	No
Q11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Q12. Were the outcome assessors blinded to the exposure status of participants?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Q13. Was loss to follow-up after baseline 20% or less?	Yes	Yes	Yes	Yes	Yes	C/D	Yes	Yes
Q14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Summary Quality	**Good**	**Good**	**Good**	**Good**	**Good**	**Good**	**Fair**	**Good**

Legends: C/D: Cannot Be Determined; N/A: Not Applicable

Table 3: Checklist for Quality Assessment for Case-control Study (Source: NIH Study Quality Assessment Tools)

Criteria	Case-control Study	
	Singh et al. (2008)	Agarwal et al. (2018)
Q1. Was the research question or objective in this paper clearly stated and appropriate?	Yes	Yes
Q2. Was the study population clearly specified and defined?	Yes	Yes
Q3. Did the authors include a sample size justification?	No	No
Q4. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?	Yes	Yes
Q5. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants?	Yes	Yes
Q6. Were the cases clearly defined and differentiated from controls?	Yes	Yes
Q7. If less than 100 percent of eligible cases and/or controls were selected for the study, were the cases and/or controls randomly selected from those eligible?	N/A	N/A
Q8. Was there use of concurrent controls?	Yes	Yes
Q9. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?	C/D	C/D
Q10. Were the measures of exposure/risk clearly defined, valid, reliable, and implemented consistently (including the same time period) across all study participants?	Yes	Yes
Q11. Were the assessors of exposure/risk blinded to the case or control status of participants?	N/A	N/A
Q12. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?	Yes	Yes
Summary Quality	**Fair**	**Fair**

Legends: C/D: Cannot Be Determined; N/A: Not Applicable

Tobacco smoking could be the most critical contributor to PAH exposure, and the exposure is not a significant health factor without it (Polanska et al., 2010). The combination of elevated DNA adducts and ETS exposure was associated with a 7% decrease in birth weight and a 3% decrease in head circumference (Perera et al., 2004).

In more recently published studies, prenatal PAHs exposure has been shown to significantly reduce birth weight, birth length, and head circumference (Jedrychowski et al., 2017; Polanska et al., 2010). In terms of impact size between PAH and $PM_{2.5}$, PAH have a greater impact size when compared to $PM_{2.5}$ (Jedrychowski et al., 2017). PAHs exposure during pregnancy causes DNA damage, leading to the induction of apoptosis, antiestrogenic effects, and binding to human aryl hydrocarbon receptors (AHR), resulting in the generation of P450 enzymes. These then interact with growth factor receptors, resulting in inefficient placentation as well as reduced oxygen and nutrient exchange to the fetus (Herbstman et al., 2012; Perera et al., 2011; Page et al., 2002).

SGA and Preterm Birth

1-In-unit increase in prenatal exposure to PAHs was significantly associated with a 2-fold increased chance of being born small for gestational age (SGA) or FGR less than 85% and an increased chance of premature delivery by five times in African Americans (Choi et al., 2008). This observation was consistent with a previous study that found SGA to be associated with PAH exposure in infants in the Czech Republic (Dejmek et al., 2000). Although the mean PAHs exposure in the sample was only one-third of that reported in the Czech Republic, the estimated risk of being born SGA among African Americans was significantly higher than that observed in the Czech population (Dejmek et al., 2000).

The level of PAHs was detected higher in placental samples of the preterm birth group (cases) compared to the full-term delivery group (controls) (Agarwal et al., 2018; Singh et al., 2008). PAHs induce the formation of reactive oxygen species (ROS) that lead to lipid peroxidation in tissues (LPO) (Al-Saleh et al., 2013; de Almeida et al., 2007). Malondialdehyde (MDA), a LPO byproduct usually regarded as a key indication of the antioxidant defence system's failure, while glutathione (GSH) acts as an efficient non-enzymatic antioxidant molecule for evaluating PAH-mediated effects (Agarwal et al., 2018). Higher placental MDA levels and lower GSH placental levels in the preterm birth group compared to the full-term delivery group indicated an antioxidant defence failure in the detoxification of excess ROS (Agarwal et al., 2018). These findings were also reported in a prior study, which revealed that PAHs induced oxidative stress, as the study detected elevated MDA levels and lower GSH levels in the blood of the bronchial asthma group (cases) compared to the control group (Suresh et al., 2009).

Accumulation of PAHs in Biological Samples

PAHs and PAH-DNA adducts were detected in the biological samples, including placental tissues, maternal blood and cord blood (Jedrychowski et al., 2017; Zhang et al., 2017; Gladen et al., 2000). Increased prenatal exposure towards airborne PAHs was significantly associated with the formation of PAH-DNA adducts in the cord blood. Fetal DNA is more susceptible to DNA damage per unit of PAH exposure because of a greater rate of formation of the adducts and/or lower ability for DNA (Jedrychowski et al., 2013). PAHs were detected in significant amounts in maternal blood, umbilical cord blood, and placental tissues (Zhang et al., 2017). This finding suggests that the accumulation of PAH-DNA adducts may not necessarily occur, even when higher levels of PAH are detected in the biological samples. Instead of inducing DNA adduct formation, PAH exposure during pregnancy may also trigger oxidative stress or bind to estrogen receptor binding sites, thereby displacing natural estrogens (Ewa & Danuta, 2017; Plíšková et al., 2002; Vondráček et al., 2002).

Vulnerable Pregnancy Windows

Increased prenatal exposure to PAHs in the first trimester was associated with the most significant decrease in fetal growth ratio, birth weight, and birth length compared to the unit impacts of PAHs in the following trimesters (Choi et al., 2012). Other epidemiologic studies have found that fetal growth rate is regulated throughout the first trimester, resulting in a progressively greater deficit as the pregnancy progresses (Milani et al., 2005; Smith, 2004; Neufeld et al., 1999). An unfavourable intrauterine environment, especially early in pregnancy, is thought to activate the fetus's survival mechanism, which protects essential organs, such as the brain and heart, while inhibiting the development of other systems (Barker, 2006).

CONCLUSION

As a conclusion, PAH exposure during pregnancy can lead to the occurrence of adverse pregnancy outcomes, including low birth weight, reduced birth length, reduced head circumference, SGA and preterm birth. However, only a few studies have shown a significant

association between exposure and the outcome. This systematic review suggests that pregnancy is a vulnerable period in life, and environmental pollutants can have adverse effects on pregnancy outcomes.

The selected articles in this systematic review only examined the association between maternal or prenatal exposure towards PAHs during pregnancy and adverse pregnancy outcomes, but no studies specifically examined the biological mechanism involved to explain the observed relationship. Several included studies only postulated possible mechanisms based on previous studies, without providing laboratory or clinical evidence to support these postulations. Future studies could design *in vitro* experimental research that integrates the biological and mechanistic investigations to provide a comprehensive understanding of maternal exposure to PAHs and their effect on pregnancy outcomes. Additionally, this study only focuses on environmental exposure to PAHs, excluding dietary consumption and occupational exposure, which may be important confounders in this observation.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to the Dar al-Hikmah Library at the International Islamic University Malaysia (IIUM) Kuantan for providing access to essential databases that facilitated the completion of this systematic review. This research was not funded by any grant. Authors declare no conflict of interest.

REFERENCES

Abdel-Shafy, H. I., & Mansour, M. S. M. (2016). A review on polycyclic aromatic hydrocarbons: Source, environmental impact, effect on human health and remediation. *Egyptian Journal of Petroleum, 25*(1), 107–123. https://doi.org/10.1016/j.ejpe.2015.03.011

Agarwal, P., Singh, L., Anand, M., & Taneja, A. (2018). Association between placental polycyclic aromatic hydrocarbons (PAHs), oxidative stress, and preterm delivery: A case–control study. *Archives of Environmental Contamination and Toxicology, 74*(2), 218–227. https://doi.org/10.1007/s00244-017-0455-0

Al-Saleh, I., Alsabbahen, A., Shinwari, N., Billedo, G., Mashhour, A., Al-Sarraj, Y., El Din Mohamed, G., & Rabbah, A. (2013). Polycyclic aromatic hydrocarbons (PAHs) as determinants of various anthropometric measures of birth outcome. *Science of the Total Environment, 444*, 565–578. https://doi.org/10.1016/j.scitotenv.2012.12.021

Anderson, L. M., Diwan, B. A., Fear, N. T., & Roman, E. (2000). Critical windows of exposure for children's health: Cancer in human epidemiological studies and neoplasms in experimental animal models. *Environmental Health Perspectives, 108*(Suppl. 3), 573–594. https://doi.org/10.1289/ehp.00108s3573

Barker, D. J. P. (2006). Adult consequences of fetal growth restriction. *Clinical Obstetrics and Gynecology, 49*(2), 270–283. https://doi.org/10.1097/00003081-200606000-00009

Chaddha, V., Viero, S., Huppertz, B., & Kingdom, J. (2004). Developmental biology of the placenta and the origins of placental insufficiency. *Seminars in Fetal and Neonatal Medicine, 9*(5), 357–369. https://doi.org/10.1016/j.siny.2004.03.006

Choi, H., Rauh, V., Garfinkel, R., Tu, Y., & Perera, F. P. (2008). Prenatal exposure to airborne polycyclic aromatic hydrocarbons and risk of intrauterine growth restriction. *Environmental Health Perspectives, 116*(5), 658–665. https://doi.org/10.1289/ehp.10958

Choi, H., Wang, L., Lin, X., Spengler, J. D., & Perera, F. P. (2012). Fetal window of vulnerability to airborne polycyclic aromatic hydrocarbons on proportional intrauterine growth restriction. *PLoS ONE, 7*(4), e35464. https://doi.org/10.1371/journal.pone.0035464

de Almeida, E. A., Bainy, A. C. D., de Melo Loureiro, A. P., et al. (2007). Oxidative stress in *Perna perna* and other bivalves as indicators of environmental stress in the Brazilian marine environment: Antioxidants, lipid peroxidation and DNA damage. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 146*(4), 588–600. https://doi.org/10.1016/j.cbpa.2006.02.040

Dejmek, J., Solanský, I., Benes, I., Lenícek, J., & Srám, R. J. (2000). The impact of polycyclic aromatic hydrocarbons and fine particles on pregnancy outcome. *Environmental Health Perspectives, 108*(12), 1159–1164. https://doi.org/10.1289/ehp.001081159

Dybing, E., Schwarze, P. E., Nafstad, P., Victorin, K., & Penning, T. M. (2013). Polycyclic aromatic hydrocarbons in ambient air and cancer. In *Air*

Pollution and Cancer (Vol. 161). International Agency for Research on Cancer Scientific Publication.

Erickson, A. C., & Arbour, L. (2014). The shared pathoetiological effects of particulate air pollution and the social environment on fetal-placental development. *Journal of Environmental and Public Health, 2014*, Article ID 901017. https://doi.org/10.1155/2014/901017

Ewa, B., & Danuta, M. Š. (2017). Polycyclic aromatic hydrocarbons and PAH-related DNA adducts. *Journal of Applied Genetics, 58*(3), 321–330. https://doi.org/10.1007/s13353-016-0380-3

Gladen, B. C., Zadorozhnaja, T. D., Chislovska, N., et al. (2000). Polycyclic aromatic hydrocarbons in placenta. *Human & Experimental Toxicology, 19*(11), 597–603. https://doi.org/10.1191/096032700671433928

Herbstman, J. B., Tang, D., Zhu, D., et al. (2012). Prenatal exposure to polycyclic aromatic hydrocarbons, benzo[a]pyrene–DNA adducts, and genomic DNA methylation in cord blood. *Environmental Health Perspectives, 120*(5), 733–738. https://doi.org/10.1289/ehp.1104056

Jedrychowski, W. A., Majewska, R., Spengler, J. D., Camann, D., Roen, E. L., & Perera, F. P. (2017). Prenatal exposure to fine particles and polycyclic aromatic hydrocarbons and birth outcomes: A two-pollutant approach. *International Archives of Occupational and Environmental Health, 90*(3), 255–264. https://pubmed.ncbi.nlm.nih.gov/28168423/

Jedrychowski, W. A., Perera, F. P., Tang, D., et al. (2013). The relationship between prenatal exposure to airborne polycyclic aromatic hydrocarbons (PAHs) and PAH–DNA adducts in cord blood. *Journal of Exposure Science & Environmental Epidemiology, 23*(4), 371–377. https://pubmed.ncbi.nlm.nih.gov/23299301/

Kim, K. H., Jahan, S. A., Kabir, E., & Brown, R. J. (2013). A review of airborne polycyclic aromatic hydrocarbons (PAHs) and their human health effects. *Environment International, 60*, 71–80. https://doi.org/10.1016/j.envint.2013.07.019

Milani, S., Bossi, A., Bertino, E., et al. (2005). Differences in size at birth are determined by differences in growth velocity during early prenatal life. *Pediatric Research, 57*(2), 205–210. https://doi.org/10.1203/01.PDR.0000148452.98518.D5

Mohd Radzi, N. A., Abu Bakar, N. K., Emenike, C. U., & Abas, M. R. (2016). Polycyclic aromatic hydrocarbons (PAHs): Contamination level and risk assessment in urban areas, Kuala Lumpur, Malaysia. *Desalination and Water Treatment, 57*(1), 171–190. https://doi.org/10.1080/19443994.2015.1021103

National Institutes of Health (NIH). (2014). *Study quality assessment tools*. U.S. Department of Health & Human Services. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

Neufeld, L., Pelletier, D. L., & Haas, J. D. (1999). The timing hypothesis and body proportionality of the intra-uterine growth retarded infant. *American Journal of Human Biology, 11*(5), 638–646. https://doi.org/10.1002/(sici)1520-6300(199909/10)11:5%3C638::aid-ajhb7%3E3.0.co;2-q

Page, T. J., O'Brien, S., Jefcoate, C. R., & Czuprynski, C. J. (2002). 7,12-Dimethylbenz[a]anthracene induces apoptosis in murine pre-B cells through a caspase-8–dependent pathway. *Molecular Pharmacology, 62*(2), 313–319. https://doi.org/10.1124/mol.62.2.313

Perera, F. (2011). Molecular epidemiology, prenatal exposure and prevention of cancer. *Environmental Health, 10*(1), 1–3. https://doi.org/10.1186/1476-069X-10-S1-S5

Perera, F. P., Tang, D., Rauh, V., Lester, K., Tsai, W.-Y., Tu, Y.-H., Weiss, L., Hoepner, L., King, J., Del Priore, G., & Lederman, S. A. (2005a). Relationships among polycyclic aromatic hydrocarbon–DNA adducts, proximity to the World Trade Center, and effects on fetal growth. *Environmental Health Perspectives, 113*(8), 1062–1067. https://doi.org/10.1289/ehp.7908

Perera, F. P., Tang, D., Whyatt, R., Lederman, S. A., & Jedrychowski, W. (2005b). DNA damage from polycyclic aromatic hydrocarbons measured by benzo[a]pyrene–DNA adducts in mothers and newborns from Northern Manhattan, the World Trade Center area, Poland, and China. *Cancer Epidemiology, Biomarkers & Prevention, 14*(3), 709–714. https://doi.org/10.1158/1055-9965.epi-04-0457

Perera, F. P., Tang, D., Tu, Y.-H., Cruz, L. A., Borjas, M., Bernert, T., & Whyatt, R. M. (2004). Biomarkers in

maternal and newborn blood indicate heightened fetal susceptibility to procarcinogenic DNA damage. *Environmental Health Perspectives, 112*(10), 1133–1136. https://doi.org/10.1289/ehp.6833

Perera, F. P., Whyatt, R. M., Jedrychowski, W., Rauh, V., Manchester, D., Santella, R. M., & Ottman, R. (1998). Recent developments in molecular epidemiology: A study of the effects of environmental polycyclic aromatic hydrocarbons on birth outcomes in Poland. *American Journal of Epidemiology, 147*(3), 309–314. https://doi.org/10.1093/oxfordjournals.aje.a009451

Plíšková, M., Vondráček, J., Vojtěšek, B., Kozubík, A., & Machala, M. (2005). Deregulation of cell proliferation by polycyclic aromatic hydrocarbons in human breast carcinoma MCF-7 cells reflects both genotoxic and nongenotoxic events. *Toxicological Sciences, 83*(2), 246–256. https://doi.org/10.1093/toxsci/kfi040

Polanska, K., Dettbarn, G., Jurewicz, J., Sobala, W., Magnus, P., Seidel, A., & Hanke, W. (2014). Effect of prenatal polycyclic aromatic hydrocarbons exposure on birth outcomes: The Polish mother and child cohort study. *BioMed Research International, 2014*, 408839. https://doi.org/10.1155/2014/408939

Polanska, K., Hanke, W., Sobala, W., & Brzeznicki, S. (2010). Exposure to polycyclic aromatic hydrocarbons and newborn biometric indicators. *International Journal of Occupational Medicine and Environmental Health, 23*(4), 339–348. https://doi.org/10.2478/v10001-010-0028-1

Ritz, B., & Wilhelm, M. (2008). Air pollution impacts on infants and children. *Southern California Environmental Report Card, 2008*, 2–6.

Singh, V. K., Singh, J., Anand, M., et al. (2008). Comparison of polycyclic aromatic hydrocarbon levels in placental tissues of Indian women with full- and preterm deliveries. *International Journal of Hygiene and Environmental Health, 211*(5–6), 639–647. https://doi.org/10.1016/j.ijheh.2007.11.004

Smith, G. C. (2004). First trimester origins of fetal growth impairment. *Seminars in Perinatology, 28*(1), 41–50. https://doi.org/10.1053/j.semperi.2003.10.012

Suresh, R., Shally, A., Mahdi, A. A., Patel, D. K., Singh, V. K., & Rita, M. (2009). Assessment of association of exposure to polycyclic aromatic hydrocarbons with bronchial asthma and oxidative stress in children: A case-control study. *Indian Journal of Occupational and Environmental Medicine, 13*(1), 33–37. https://doi.org/10.4103/0019-5278.50722

Tang, D., Li, T. Y., Liu, J. J., Chen, Y. H., Qu, L., & Perera, F. (2006). PAH–DNA adducts in cord blood and fetal and child development in a Chinese cohort. *Environmental Health Perspectives, 114*(8), 1297–1300. https://doi.org/10.1289/ehp.8939

Vondráček, J., Kozubík, A., & Machala, M. (2002). Modulation of estrogen receptor-dependent reporter construct activation and G_0/G_1–S-phase transition by polycyclic aromatic hydrocarbons in human breast carcinoma MCF-7 cells. *Toxicological Sciences, 70*(2), 193–201. https://doi.org/10.1093/toxsci/70.2.193

Wang, L., & Choi, H. (2014). Using semiparametric-mixed model and functional linear model to detect vulnerable prenatal window to carcinogenic polycyclic aromatic hydrocarbons on fetal growth. *Biometrical Journal, 56*(2), 243–255. https://doi.org/10.1002/bimj.201200132

Zhang, X., Li, X., Jing, Y., Fang, X., Zhang, X., Lei, B., & Yu, Y. (2017). Transplacental transfer of polycyclic aromatic hydrocarbons in paired samples of maternal serum, umbilical cord serum, and placenta in Shanghai, China. *Environmental Pollution, 222*, 267–275. https://doi.org/10.1016/j.envpol.2016.12.046

ORIGINAL ARTICLE

Comparison of the Radioactivity Level of Bird's Eye Chillies in Chemical and Natural Fertilizers

Fatin Syakirah Binti Mohd Salleh[1], Fatihah Syafinaz Binti Kamarul Zaman[1,*], Norhanna Binti Sohaimi[1]

[1]Department of Diagnostic Imaging and Radiotherapy, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan Pahang, Malaysia

Keywords:
radioactivity in plants; radionuclide level; radiological hazard; radioactivity in agriculture

ABSTRACT

Background: Concerns regarding naturally occurring radionuclides in agricultural inputs have increased due to their potential transfer into food crops. This study evaluates the activity concentrations of K-40, Th-232, and Ra-226 in chemical and natural fertilizers and examines their subsequent uptake in bird's eye chillie**s** to assess radiological safety. **Methods:** Twelve samples comprising chemical fertilizers, natural fertilizers and bird's eye chillies were collected and analyzed using a High Sensitivity Gamma Spectrometer GDM 10C equipped with WinDAS software. Activity concentrations of the selected radionuclides were quantified, and further assessments of radium equivalent activity (Ra_{eq}), absorbed dose rate, and annual effective dose were performed. Data processing and comparative analyses were conducted using Microsoft Excel. **Results:** All three radionuclides were detected in both fertilizer types. Activity concentrations in fertilizers ranged from 1.3329×10^{-5} to 6.3679×10^{-5} Bq/kg, while bird's eye chillies samples showed higher values ranging from 3.9453×10^{-5} to 9.7137×10^{-5} Bq/kg. Chemical fertilizers exhibited higher radioactivity indicators, with average Ra_{eq} (3.6147 Bq/kg), absorbed dose rate (1.6472 nGy/h), and annual effective dose (0.0202 mSv) compared to natural fertilizers (1.7146 Bq/kg, 0.7941 nGy/h, and 0.0097 mSv, respectively). Correspondingly, chillies grown with chemical fertilizers showed a higher annual ingestion dose of K-40, Th-232, and Ra-226. **Conclusion**: Although radioactivity levels were higher in samples influenced by chemical fertilizers compared to natural fertilizers, all measured doses remained well below the public exposure limit of 1 mSv/year. Therefore, the bird's eye chillies samples analyzed in this study are considered radiologically safe for human consumption.

INTRODUCTION

Radioactivity is a natural phenomenon caused by the spontaneous decay of unstable atomic nuclei, emitting various forms of radiation such as alpha, beta, and gamma rays. Radioactive elements are present in varying concentrations and can be found in natural and man-made substances. Among these substances are chemical and natural fertilizers, which play a crucial role in modern agriculture by providing essential nutrients for plants. However, concerns have arisen regarding the potential presence of radioactive elements in fertilizers and their influence on the activity concentration in plants.

According to Boukhenfouf & Boucenna (2011), soil nutrients are replaced by chemical fertilizers, primarily compounds marketed as NPKs (nitrogen (N), phosphorus (P), and potassium (K)), which provide the necessary elements to attain high agricultural output. Phosphorus is sourced from phosphate rocks, which contain elevated levels of naturally occurring Uranium-238, Thorium-232, and K-40 and their radioactive daughters. There is a significant risk towards the consumers, workers, members of the public, and the environment regarding the radiation hazards of fertilizer materials as it contains naturally occurring radioactive materials (Sarkar et al., 2024).

A previous study by Nguyen et al. (2020) examined the amount of radioactivity transferred from soil to plants. However, since the study focused on high-risk radiation areas, it did not account for radioactivity levels in fertilizers. In other study that measured the radioactivity level in fertilizer, it did not focus on the effect of the edible plants (Thabayneh & Hreaz, 2025). Therefore, this study is important because understanding radionuclide transfer and accumulation in crops is essential for food safety, environmental protection, and public health. Even when measured activity concentrations may fall below the safety threshold of 1 mSv/year, chronic low-level exposure from continuous dietary intake may still pose long-term health risks, especially in regions where chemical fertilizers are widely used.

This study aimed to provide an overview of the radioactivity levels of Ra-226, Th-232, and K-40 in chemical

* *Corresponding author.*
E-mail address: syafinaz@iium.edu.my

and natural fertilizers and to compare their influence on the radioactivity level in bird's eye chilies to ensure it is safe and harmless for consumers to ingest. Specifically, the study evaluated whether fertilizer type affects the radionuclide uptake in the edible portion of the plant and whether the resulting levels remain within radiological safety limits for consumers, while providing baseline data of the natural radioactivity level of bird's eye chilies to inform safer agricultural practices.

MATERIALS AND METHODS

Material and Equipment

Radionuclide activities in samples of chemical fertilizer, natural fertilizer and bird's eye chili were quantified using a High Sensitivity Gamma Spectrometer (Model GDM 10C) as in Figure 1 coupled with WinDAS software. The GDM 10C is a scintillation detector equipped with a 2x2 inch sodium iodide (NaI) crystal. The detector assembly, measuring 12 cm in diameter and 40 cm in height, is encased in 1-inch-thick cylindrical lead shielding to minimize background radiation.



Figure 1: High Sensitivity Gamma Spectrometer GDM 10C

The WinDAS software version 3.4.20 was used to measure and analyze gamma radiation from the radionuclide activity in the samples. The software is designed to automatically determine several radionuclides available in the data library such as Ra-226, Th-232, and K-40. The frequency diagram of the energy distribution of the detected gamma quanta was demonstrated in Figure 2.

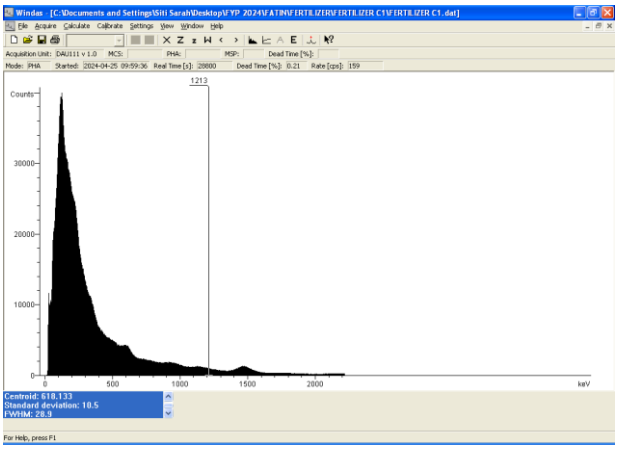


Figure 2: WinDAS Software

Methodology

Sample selection

Radionuclide activities in samples of chemical fertilizer, natural fertilizer, and bird's eye chili were quantified using a High Sensitivity Gamma Spectrometer (Model GDM 10C) coupled with WinDAS software. The GDM 10C is a scintillation detector equipped with a 2x2 inch sodium iodide (NaI) crystal. The detector assembly, measuring 12 cm in diameter and 40 cm in height, is encased in 1-inch thick cylindrical lead shielding to minimize background radiation.

The experimental framework involved the cultivation of ten (n=10) bird's eye chillies saplings to assess the transfer of radionuclides. The saplings were potted individually in a controlled domestic environment to simulate typical small-scale agricultural conditions. To evaluate the impact of different nutrient sources, the plants were segregated into two distinct experimental cohorts, consisting of five replicates each (n=5).

The first cohort was subjected to a regimen of commercial chemical fertilizer, while the second cohort was treated exclusively with natural organic fertilizer. All other environmental variables, including sunlight exposure and irrigation, were maintained consistently across both groups. Upon reaching maturity, representative samples were systematically collected for radiological analysis. The final sample inventory comprised four distinct matrices: (1) raw chemical fertilizer, (2) raw natural fertilizer, (3) chili harvested from the chemically treated saplings, and (4) chili harvested from the organically or naturally treated saplings. The visual for the sample selection is shown in Figure 3.

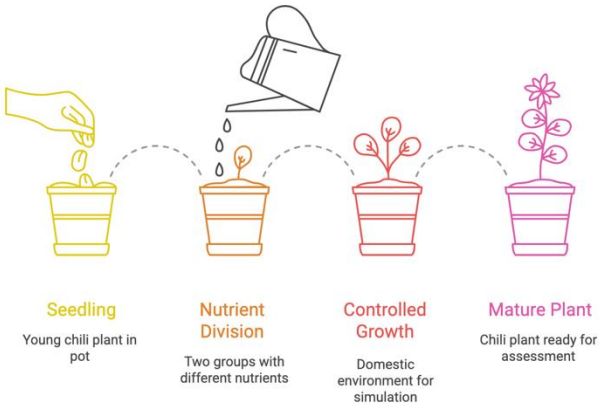


Figure 3: Visual for Chili Sapling

Sample preparation

A total of 12 samples were prepared, comprising three chemical fertilizer samples, three natural fertilizer samples, three bird's eye chillies samples cultivated with chemical fertilizer, and three bird's eye chillies samples cultivated with natural fertilizer. The fertilizer samples were dried in a dehydrator at 55°C for 8 h to remove moisture.

For the chilli samples, they were rinsed with tap water to remove adhering dirt then dried in a dehydrator at 55°C for 12 h to ensure moisture elimination. Longer time required to eliminate the moisture for chilli samples. Subsequently, all samples were ground to a fine powder using a mortar and pestle to achieve homogenization. Each powdered sample was weighed to 10 g on a digital balance, placed in labelled containers as listed in Table 1 and stored at room temperature for 28 days to allow radioactive equilibrium to be established between parent nuclides and their daughter products prior to gamma-ray spectrometric analysis (Kessaratikoon et al., 2021).

Table 1: Samples and its coded label

Type of sample	Coded label
Natural fertilizer	NF1
*NF = natural fertilizer	NF2
	NF3
Chemical fertilizer	CF1
*CF = chemical fertilizer	CF2
	CF3
Bird's eye chillies fertilized by natural fertilizer	NC1
*NC = natural chilli	NC2
	NC3
Bird's eye chillies fertilized by chemical fertilizer	CC1
*CC = chemical chilli	CC2
	CC3

Sample reading

The radionuclide activities were determined using the GDM 10C gamma spectrometer equipped with a sodium iodide (NaI) detector. The system was calibrated with a Co-60 source to quantify activity levels and to convert channel numbers on the x-axis to corresponding energies expressed in keV. Sample measurements for chemical and natural fertilizers were conducted for 28,800 seconds (8 hours), while bird's eye chillies samples were measured for 18,000 seconds (5hours) to minimise statistical uncertainty. Background radiation was recorded by operating the detector without any sample. Net counts for each sample were obtained by subtracting the background spectrum from the respective sample spectra. Each sample was measured in triplicate to derive mean activity values. The flow of the sample selection and sample preparation were shown in Figure 4. The counts corresponding to the characteristic energy peaks of each radionuclide are presented in Table 2.

Table 2: Source, energy, and emission probability of the radionuclides

Source	Daughter Nuclide	Energy (keV)	Emission Probability (%)
K-40	-	1460.8	10.7
Th-232	Ac-228	911.1	29.0
	Tl-208	583.1	86.0
Ra-226	Pb-214	352.0	37.1
	Bi-214	609.3	46.1

Sample analysis

All data were analyzed using Microsoft Excel. The specific activity concentration, A (Bq/kg) of each radionuclide in fertilizers and bird's eye chillies samples were calculated using Equation 1 as stated in a study conducted by Salih et al. (2019) :

$$A, (Bq.kg^{-1}) = ((Net\ count))/(\varepsilon \times I\gamma \times t \times m) \quad (1)$$

where, Iɣ = emission probability per decay of the specific peak
ε = absolute gamma peak efficiency for the detector at a particular photopeak
t = is the counting time in seconds
m = mass of the sample in kilogram.

To assess the real activity level of K-40, Th-232, and Ra-226 in fertilizers, a common radiological index has been

defined in terms of radium equivalent activity (Ra_{eq}) in Bq/kg can be used, providing a very useful guideline in regulating the safety standards in radiation protection for a human population. The index was calculated through the following formula by Alharbi (2013) :

$$Ra_{eq}\ (\text{Bq/kg}) = C_{Ra} + 1.43C_{Th} + 0.077C_K \qquad (2)$$

where, C_{Ra} = specific activities (Bq/kg) of Ra-226
C_{Th} = specific activities (Bq/kg) of Th-232
C_K = specific activities (Bq/kg) of K- 40

The measured activity of K-40, Th-232, and Ra-226 were converted into doses by applying the factors 0.0417, 0.604, and 0.462 respectively (UNSCEAR, 2000). These factors were used to calculate the total absorbed gamma dose rate in air at 1 m above the ground level using the following equation by Alharbi (2013):

$$D\ (nGy/h) = 0.462C_{Ra} + 0.604C_{Th} + 0.0417C_K \quad (3)$$

where, C_{Ra} = specific activities (Bq/kg) of Ra-226
C_{Th} = specific activities (Bq/kg) of Th-232
C_K = specific activities (Bq/kg) of K- 40

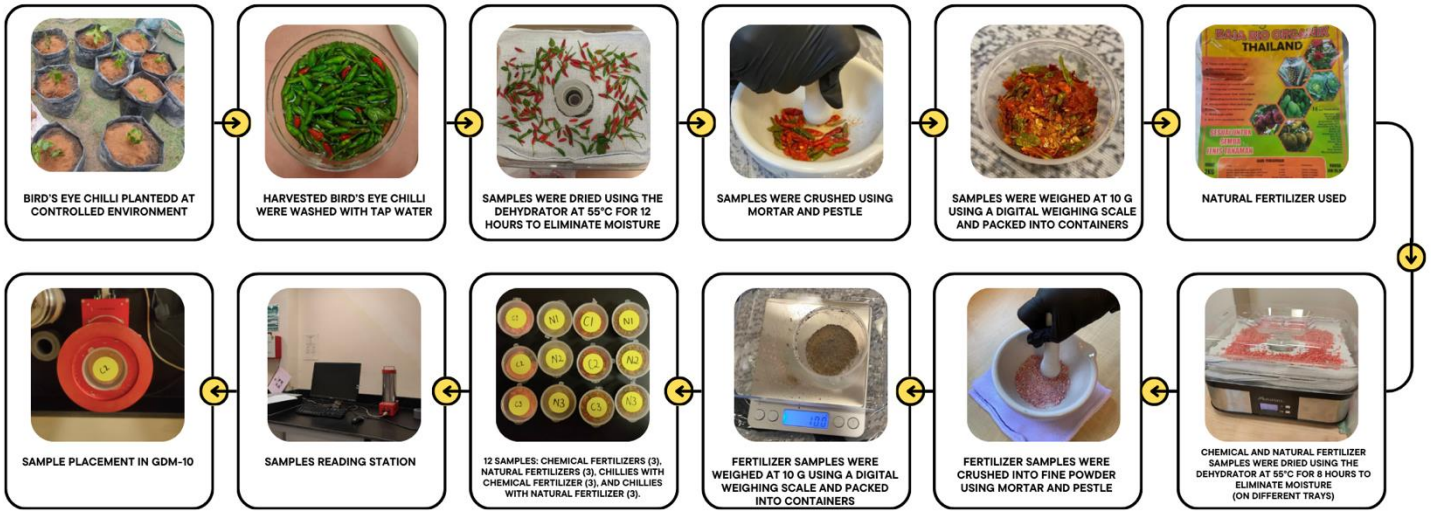


Figure 4: Flow of sample selection and sample preparation

To estimate annual effective doses, must be considered of the conversion coefficient from absorbed dose in air to effective dose. The annual effective doses are determined as follows:

$$AED\ (mSv) = D \text{ x } 8760 \text{ h x } 0.2 \text{ x } 0.7\,\frac{Sv}{Gy}\ x\ 10^{-6} \quad (4)$$

where, D = absorbed dose
8760 = number of hours in a year
0.2 = outdoor occupancy factor

Annual estimated average effective dose equivalent received by a member is calculated using a conversion factor of 0.7 Sv/Gy.

RESULTS

Radioactivity Level in Fertilizers

The measured activity concentration of K-40, Th-232, and Ra-226 in chemical and natural fertilizers is presented in Table 3. Each sample of both chemical and natural fertilizers have different activity concentrations of K-40, Th-232, and Ra-226. However, there are also the studied radionuclides not detected in several samples.

Table 3: Activity concentration in chemical fertilizer and natural fertilizer

Sample	Activity concentration (×10^{-5} Bq/kg)		
	K-40	Th-232	Ra-226
NF1	1.3773	0.0459	ND
NF2	ND	ND	4.9720
NF3	ND	ND	ND
Average	0.4591	0.0153	1.6573
CF1	1.5994	0.3088	ND
CF2	1.3329	1.1738	2.0739
CF3	ND	0.0394	6.3679
Average	0.9774	0.5073	2.8139
Minimum	1.3329	0.0394	2.0739
Maximum	1.5994	1.1738	6.3679
Mean	0.7182	0.2613	2.2356

The activity concentration in all fertilizer samples is in the range of 1.3329×10^{-5} Bq/kg to 1.5994×10^{-5} Bq/kg for K-40, 0.0394×10^{-5} Bq/kg to 1.1738×10^{-5} Bq/kg for Th-232, and 2.0739×10^{-5} Bq/kg to 6.3679×10^{-5} Bq/kg for Ra-226. The highest and lowest activity of K-40, Th-232, and Ra-226 was found in chemical fertilizer samples.

Radioactivity Level in Bird's Eye Chili

The measured activity concentration of K-40, Th-232, and Ra-226 in bird's eyes chillies fertilized by chemical and natural fertilizers is presented in Table 4. All samples of bird's eye chillies have different activity concentrations of K-40, Th-232, and Ra-226. The activity concentration in all bird's eye chilli samples is in the range of 3.9453×10^{-5} Bq/kg to 6.3622×10^{-5} Bq/kg for K-40, 1.3382×10^{-5} Bq/kg to 3.0362×10^{-5} Bq/kg for Th-232, and 1.7504×10^{-5} Bq/kg to 9.7137×10^{-5} Bq/kg for Ra-226.

Table 4: Activity concentration in bird's eye chillies fertilized with chemical and natural fertilizer

Sample	Activity concentration ($\times10^{-5}$ Bq/kg)		
	K-40	Th-232	Ra-226
CC1	6.3622	3.0362	2.3943
CC2	5.8646	2.9241	9.0190
CC3	4.7983	1.8308	9.7137
Average	5.6750	2.5970	7.0423
NC1	3.9453	2.5180	6.0198
NC2	4.1230	1.3382	2.4519
NC3	5.5092	2.7858	1.7504
Average	4.5258	2.2140	3.4073
Minimum	3.9453	1.3382	1.7504
Maximum	6.3622	3.0362	9.7137
Mean	5.1004	2.4055	5.2249

The highest activity concentration of K-40, Th-232, and Ra-226 was found in bird's eye chillies fertilized with chemical fertilizer samples. The lowest activity concentration of K-40, Th-232, and Ra-226 was found in bird's eye chillies fertilized with natural fertilizer samples. Figure 5 compares the annual ingestion doses of K-40, Th-232, and Ra-226 across all samples. Overall, the comparison indicates that chemical fertilizer is associated with higher radionuclide ingestion doses, particularly for Ra-226.

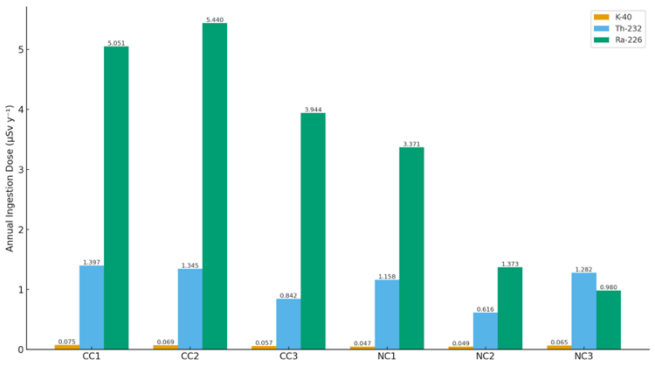


Figure 5: Comparison of the annual ingestion doses of K-40, Th-232, and Ra-226 across all samples

Radiological Health Hazard

The Radium equivalent activity, absorbed dose, and annual effective dose were calculated using the specific activity concentration of the radionuclides in each fertilizer sample as shown in Table 5. Radium equivalent is used to assess the potential health risks associated with natural gamma emitters present in NORM, as it allows to unify the radiation exposure effect caused by different isotopes. Chemical fertilizer has higher Radium equivalent activity with the average of 3.6147 Bq/kg compared to natural fertilizer which only gives the average reading of 1.7146 Bq/kg.

Table 5: Radium equivalent activity, absorbed dose, and annual effective dose of K-40, Th-232, and Ra-226 for chemical and natural fertilizers

Sample	Ra_{eq} (Bq/kg)	Absorbed dose (nGy/h)	Annual effective dose (mSv)
NF1	0.1717	0.0852	0.0010
NF2	4.9720	2.2971	0.0282
NF3	-	-	-
Average	1.7146	0.7941	0.0097
CF1	0.5647	0.2532	0.0031
CF2	3.8551	1.7227	0.0211
CF3	6.4242	2.9658	0.0364
Average	3.6147	1.6472	0.0202

Absorbed dose is a vital factor in radiation protection as it allows estimating the potential negative effect of radiation from soil on human. Chemical fertilizer has higher absorbed dose with the average of 1.6472 nGy/h compared to natural fertilizer which only gives the average reading of 0.7941 nGy/h. Annual effective dose can be obtained by using the absorbed dose. Chemical fertilizer

shows a higher value of annual effective dose compared to natural fertilizer with the average of 0.0202 mSv and 0.0097 mSv respectively.

The annual ingestion dose of K-40, Th-232, and Ra-226 for bird's eye chillies fertilized with chemical and natural fertilizers are presented in Table 6. The dose due to ingestion of Ra-226 has a maximum value of 5.0506 μSv y^{-1} in the bird's eye chillies fertilized with chemical fertilizer. The minimum dose was detected in the Bird's Eye chilli fertilized with natural fertilizer of 0.9802 μSv y^{-1} for Ra-226.

Table 6: Annual ingestion dose of K-40, Th-232, and Ra-226 for Bird's Eye chilli fertilized with chemical and natural fertilizers

Sample	Annual Ingestion Dose (μSv y^{-1})		
	K-40	**Th-232**	**Ra-226**
CC1	0.0751	1.3967	5.0506
CC2	0.0692	1.3451	5.4397
CC3	0.0566	0.8422	3.9437
Average	0.0670	1.1946	4.8113
NC1	0.0466	1.1583	3.3711
NC2	0.0487	0.6156	1.3731
NC3	0.0650	1.2815	0.9802
Average	0.0534	1.0184	1.9081
Minimum	0.0466	0.6156	0.9802
Maximum	0.0751	1.3967	5.0506
Range	0.0466 – 0.0751	0.6156 – 1.3967	0.9802 – 5.0506

The dose due to Th-232 has a range between 0.6156 μSv y^{-1} for bird's eye chillies fertilized with natural fertilizer and 1.3967 μSv y^{-1} for bird's eye chillies fertilized with chemical fertilizer. Bird's eye chillies fertilized with chemical fertilizer recorded the maximum K-40 dose of about 0.0751 μSv y^{-1}, while the minimum dose was found in Bird's Eye chilli fertilized with natural fertilizer of about 0.0466 μSv y^{-1}.

DISCUSSION

This study examined the activity concentrations of K-40, Th-232, and Ra-226 in chemical and natural fertilizers and their transfer to bird's eye chillies plants. Results showed that Ra-226 exhibited the highest activity concentration in all samples, followed by K-40, while Th-232 recorded the lowest. Chemical fertilizers consistently showed higher radionuclide concentrations than natural fertilizers.

The elevated levels in chemical fertilizers are mainly attributed to the phosphate rocks used in their production, which naturally contain uranium and thorium series radionuclides that persist in the final product (Ahmad et al., 2015). Similarly, the higher K-40 concentration is linked to potash, a potassium-rich compound used in many chemical fertilizers that inherently contains radioactive isotopes of potassium (Ogundele et al., 2026). In contrast, natural fertilizers derived from organic matter and animal manure showed much lower radionuclide activity.

Other than that, bird's eye chillies plants grown with chemical fertilizers displayed higher activity concentrations of K-40, Th-232, and Ra-226 than those treated with natural fertilizers, with Ra-226 being dominant. This enrichment reflects the transfer of radionuclides from fertilizers to plants, consistent with findings that phosphate-based fertilizers enhance radionuclide uptake due to their inherent radioactivity (Mankala et al., 2025).

The mobility and bioavailability of radionuclides in soil play a crucial role in their uptake. Chemical fertilizers can alter soil pH and introduce complexing agents that increase radionuclide solubility that facilitating root absorption. The elevated K-40 concentrations are further explained by potassium's essential role in plant nutrition and its uptake regardless of isotope form.

These findings have important environmental and health implications. Continuous application of phosphate-based fertilizers may lead to radionuclide accumulation in soil and their subsequent transfer to crops, increasing radiation exposure through the food chain (Sahu et al., 2014).

CONCLUSION

This study aims to measure and compare the radioactivity level of K-40, Th- 232, and Ra-226 in chemical fertilizer and natural fertilizer and their influence on the radioactivity level of bird's eye chillies. This experimental study was conducted using High Sensitivity Gamma Spectrometer GDM 10C. The findings reveal that the activity concentration of K-40, Th-232, and Ra-226 is higher in chemical fertilizer and Bird's Eye chilli treated with chemical fertilizer than natural fertilizer and Bird's Eye chilli fertilized with natural fertilizer.

For radiological health hazard, the Radium equivalent activity, absorbed dose rate, and annual effective dose of K-40, Th-232, and Ra-226 for chemical fertilizers gives higher values than natural fertilizers. On the other hand, the annual ingestion dose of K-40, Th-232, and Ra-226 for bird's eye chillies fertilized with chemical fertilizers is

higher than fertilized with natural fertilizers. However, the values are still considered safe.

ACKNOWLEDGEMENT

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

Ahmad, N., Jaafar, M. S., & Alsaffar, M. S. (2015). Natural radioactivity in virgin and agricultural soil and its environmental implications in Sungai Petani, Kedah, Malaysia. *Polymer Journal*, 1, 305–313. https://api.semanticscholar.org/CorpusID:54742555

Alharbi, W. (2013). Natural Radioactivity and Dose Assessment for Brands of Chemical and Organic Fertilizers Used in Saudi Arabia. *Journal of Modern Physics*, 4, 344–348. https://doi.org/10.4236/jmp.2013.43047

Boukhenfouf, W., & Boucenna, A. (2011). The radioactivity measurements in soils and fertilizers using gamma spectrometry technique. *Journal of Environmental Radioactivity*, 102(4), 336–339. https://doi.org/https://doi.org/10.1016/j.jenvrad.2011.01.006

Kessaratikoon, P., Riyapunt, D., Boonkrongcheep, R., & Changkit, N. (2021). Assessment of contamination of natural and anthropogenic radionuclides in rice samples collected from Songkhla province, Thailand. *Journal of Physics: Conference Series,* 2145, 12019. https://doi.org/10.1088/1742-6596/2145/1/012019

Mankala, J. B., Sawe, S. F., & Moirana, R. L. (2025). Assessing the impact of phosphate fertilizer application on radionuclide accumulation in soil and Spinacia oleracea. *Journal of Ecological Engineering,* 26(9), 38–48. https://doi.org/10.12911/22998993/204696

Nguyen, V. T., Thu Huynh, N. P., & Le, C. H. (2020). Accumulation rates of natural radionuclides (40K, 210Pb, 226Ra, 238U, and 232Th) in topsoils due to long-term cultivations of water spinach (Ipomoea Aquatica Forssk.) and rice (Oryza Sativa L.) based on model assessments: A case study in Dong Nai province, Vietnam. *Journal of Environmental Management,* 271, 111001. https://doi.org/10.1016/j.jenvman.2020.111001

Ogundele, L. T., Orosun, M. M., Akiniran, S., & Ayeku, P. O. (2026). Activity levels and radiological hazards of chemical fertilizers used for farm crops in Ondo city, Southwest, Nigeria. *Nuclear Analysis*, 5(1), 100201. https://doi.org/10.1016/j.nucana.2025.100201

Sahu, S. K., Ajmal, P. Y., Bhangare, R. C., Tiwari, M., & Pandit, G. G. (2014). Natural radioactivity assessment of a phosphate fertilizer plant area. Journal of Radiation *Research and Applied Sciences,* 7(1), 123–128. https://doi.org/10.1016/j.jrras.2014.01.001

Salih, N. F., Hussein, Z. A., & Sedeeq, S. Z. (2019). Environmental radioactivity levels in agricultural soil and wheat grains collected from wheat-farming lands of Koya district, Kurdistan region-Iraq. *Radiation Protection and Environment*, 42(4). https://journals.lww.com/rpae/fulltext/2019/42040/environmental_radioactivity_levels_in_agricultural.2.aspx

Sarkar, K., Chakraborty, A. K., Ahmed, M., & Hossain, S. (2024). Assessment of naturally occurred radiation hazards in chemical and organic fertilizers used in Chattogram area, Bangladesh. *International Journal of Environmental Analytical Chemistry,* 104(17), 5986–5998. https://doi.org/10.1080/03067319.2022.2135998

Thabayneh, K. M., & Hreaz, H. M. (2025). Measurement of radioactive contamination of plant fertilizers used in Palestine. *Nuclear Engineering and Technology,* 57(9), 103607. https://doi.org/10.1016/j.net.2025.103607

UNSCEAR. (2000). Sources and Effects of Ionizing Radiation. *United Nation Publication.* https://www.unscear.org/unscear/uploads/documents/unscear-reports/UNSCEAR_2000_Report_Vol.II.pdf

ORIGINAL ARTICLE

Prevalence and Ergonomic Risk Profiling of Musculoskeletal Pain in Medical Imaging Students: A Basis for Prevention and Intervention.

Hanna Yasmin Harun[1], Liyana Musa[1,*]

[1]Department of Diagnostic Imaging and Radiotherapy, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
musculoskeletal pain; medical imaging students; ergonomics risk factors; clinical task

ABSTRACT

Background: Musculoskeletal pain (MSP) is a common problem among university students, particularly those in health-related programmes with significant physical and ergonomic demands during clinical training, and medical imaging students are likely to be similarly affected. The objectives of this study were to identify the risk factors associated with MSP, evaluate its prevalence, frequency, and severity, as well as examine the correlation between MSP frequency and potential risk factors among medical imaging students. **Methods:** A cross-sectional study was conducted using a validated questionnaire. Descriptive statistics and Spearman's correlation were used to analyse the correlation between MSP and potential risk factors. **Results:** The results showed that the most frequently affected body regions were the shoulders (49.2%), lower back (38.1%), and neck (36.5%), with both acute (short-term) and persistent (long-term) pain reported. MSP frequency (past 12 months) showed significant correlations with clinical-task-related variables, particularly repetitive motions (ρ=0.787), extended standing (ρ=0.683), and incorrect postures (ρ=0.744). In contrast, lifestyle factors such as sleep duration and stress management showed weaker associations. **Conclusion:** These findings highlight the need for ergonomic training and preventive strategies to reduce the risk of MSP among medical imaging students. Future studies should examine long-term effects and evaluate intervention strategies to improve clinical performance and student well-being.

INTRODUCTION

Musculoskeletal pain (MSP) is a common disorder affecting bones, muscles, ligaments, tendons, and nerves, and may present as acute or chronic pain. Acute MSP is localised to a specific body region (e.g. back, neck, shoulder) and is usually related to structures in that area. Chronic MSP refers to pain that persists for more than three months and may be linked to illness or injury, influenced by psychological factors, lack clear biological function, and have no identifiable endpoint.

Pain related to MSP is a frequent medical and socioeconomic problem globally (El-Tallawy et al., 2021). MSP can cause physical discomfort, limit daily activities, and reduce overall quality of life. It can also affect mental health by increasing stress and anxiety and reducing productivity.

The impact of MSP extends beyond individual health and can disrupt healthcare students' engagement in practical training and clinical placements. High MSP prevalence has been reported among dental, healthcare, and allied health students (91.2%, 64.8%, and 73.6%, respectively) (Abdulaziz et al., 2019; Felemban et al., 2021; Senarath et al., 2021; Syafiqah et al., 2021). However, medical imaging students have received comparatively less attention, despite facing similar challenges in their learning and clinical environments.

Medical imaging students are often required to stand for long periods, perform repetitive movements, and maintain awkward postures during imaging procedures. Combined with heavy equipment handling and precise patient positioning, these demands increase their susceptibility to MSP. Extended standing, for example, can lead to foot pain and swelling, varicose veins, fatigue, stiff neck and shoulders, and low back pain (Alias et al., 2020). Yet ergonomic training and interventions for this group are often limited or overlooked, potentially exacerbating these risks.

Although MSP has been widely reported among healthcare and allied health students, evidence on the detailed pattern of MSP (location, severity and frequency) and its ergonomic/lifestyle risk factors specifically in medical imaging students, especially in Malaysia remains limited. Studies involving students in medical imaging or radiography programmes (e.g, Italy, South Africa) have mainly focused on prevalence and a few occupational

* *Corresponding author.*
E-mail address: liyanamusa@iium.edu.my

factors, often using study-specific or adapted questionnaires with limited psychometric reporting, despite the availability of validated tools such as the Standardised Nordic Musculoskeletal Questionnaire (Lee & Cho, 2012; Lorusso et al., 2010; Masondo & Khoza, 2023; Syafiqah et al., 2021).

This study adds value by using a validated MSP instrument with 12-month and 7-day recall and a structured risk-factor section to quantify MSP prevalence, frequency, and severity across body regions and examine correlations between the frequency of MSP in the past 12 month with the potential risk factors associated with MSP among medical imaging students. The findings provide evidence to guide early detection and targeted prevention strategies such as ergonomic education, workload modification, and exercise programmes to protect students' musculoskeletal health and reduce future work-related problems.

MATERIALS AND METHODS

Study Design and Setting

This cross-sectional study specifically included medical imaging students who had undergone clinical training to ensure better representation of the experiences related to MSP in clinical training. The study was conducted from March 2024 until February 2025 with the approval of the institutional research and ethics committee.

Population and sampling

The study recruited medical imaging students from Faculty of Allied Health Sciences from a university in Pahang, Malaysia. Applying Slovin's formula at a 5% margin of error (95% confidence interval) to the population of 74 medical imaging students (Years 2-4) produced a minimum sample size of 63, which was then recruited using convenience sampling.

Questionnaires

Data were collected using a three-part questionnaire: Section 1 covered demographics, Section 2 assessed MSP prevalence and characteristics, and Section 3 examined potential risk factors. Risk-factor items described task-related and postural exposures (e.g. prolonged neck flexion/extension, patient lifting) treated as potential MSP risk factors.

The questionnaire was designed using a five-point Likert-scale which consists of a series of statements response to indicate respondents' pain severity, frequency, and the potential risk factors. These questions were also constructed using an adopt-and-adapt approach based on established and validated MSP instruments, particularly the Standardised Nordic Musculoskeletal Questionnaire by Kuorinka et al. (1987). The prevalence section follows the Nordic convention by assessing symptoms across two recall periods: the past 12 months and the past 7 days.

The 12-month timeframe captures annual prevalence and reflects longer-term or recurrent MSP patterns that may occur intermittently over an academic year. While the 7-day timeframe captures recent prevalence, providing a snapshot of current or very recent symptoms and helping to minimise recall bias. Using both periods enables a clearer distinction between longer-term or recurrent and recent MSP experiences among the students (Kuorinka et al., 1987; Dawson et al., 2009; Legault et al., 2014; Rasmussen et al., 2018; Umar et al., 2025).

The Likert-scale questionnaire was reviewed by five experts (medical imaging academicians and practising radiographers). The validation ensured that items were relevant, clear, and appropriate for medica lmaging students, and minimised leading or biased wording. Based on their feedback on redundancy, ambiguity, and consistency, the questionnaire was refined before distribution.

Cronbach's alpha was used to assess internal consistency. Alpha coefficients were 0.743 for MSP experience, 0.891 for severity, 0.909 for frequency, and 0.799 for potential risk factors, all above the 0.70 threshold (Bonett & Wright, 2015). This indicates acceptable to excellent reliability, suggesting the questionnaire is suitable for assessing MSP prevalence, characteristics, and related risk factors among medical imaging students.

Data Collection

All respondents received a brief study information sheet explaining the study purpose, and confidentiality of their responses before given the questionnaire. Participation was entirely voluntary; informed consent was obtained prior to data collection, and they could withdraw at any point without any penalty or loss of benefits. The online questionnaire (via Google Forms) was open for one month, with reminders sent three times per week to medical imaging students, and this period was deemed sufficient for data collection.

Analysis

All completed responses were analysed using the Statistical Package for the Social Sciences (SPSS) version 20. Demographic characteristics, MSP prevalence and severity, and potential risk factors were summarised using

descriptive statistics (frequencies, percentages, means, and standard deviations).

The study also investigated whether there was any correlation between the frequency of MSP in the past 12 months and the identified potential risk factors. The 12-month measure was used because it reflects longer-term or recurrent symptoms and is more consistent with habitual ergonomic and lifestyle exposures. The 7-day measure was reported descriptively only, as it represents short-term, potentially transient symptoms.

Before correlation analysis, data normality was assessed using the Kolmogorov–Smirnov test ($p > 0.05$ indicating normal distribution). Correlations were considered statistically significant at $p \leq 0.05$, and their strength was interpreted from the correlation coefficient (r), with values closer to 0 indicating negligible association and larger absolute values (approximately 0.1–0.3 small, 0.3–0.5 moderate, >0.5 strong) indicating progressively stronger associations (Laerd Statistics, 2018).

RESULTS

Demographic Characteristics of Respondents

A total of 63 students pursuing degrees in medical imaging participated in the study (16 males, 47 females) from Years 2–4 (Year 2: 18, Year 3: 25, Year 4: 20). Over half had a BMI in the normal range (n = 37, 58.7%), with the remainder distributed across underweight (n = 11), overweight (n = 9), and obese or extremely obese categories (n = 6). Only four students (6.3%) reported a previous history of MSP. Fewer than one-quarter (n = 15, 23.8%) reported using ergonomic equipment, and none reported smoking or vaping.

Site-specific Prevalence of MSP

This section reports the site-specific prevalence of MSP shown as number(n) and percentage (%) of students who reported MSP in each body region at least once during the specified period. Students were permitted to choose various sites, resulting in percentages that do not sum to 100%.

For the duration of past 12 months (Table 1), the most frequently cited sites were the shoulders (n = 31, 49.2%), lower back (n = 24, 38.1%), neck (n = 23, 36.5%), and ankles/feet (n = 20, 31.7%). Moderate numbers of students also reported pain in the upper back and hips/thighs (each n = 17, 27.0%), followed by the knee (n = 14, 22.2%) and wrist/hands (n = 13, 20.6%), while the elbow was least reported (n = 6, 9.5%).

Meanwhile for the duration of past 7 days (Table 1), lower back pain was the most frequently reported (n = 16, 25.4%), followed by pain in the neck (n = 10, 15.9%) and knees (n = 9, 14.3%). Shoulder and wrist/hand pain were each reported by eight students (12.7%), whereas pain in the upper back, hips/thighs and ankles/feet was less common (each n = 3, 4.8%). No students reported elbow pain in the preceding week.

Table 1: Experience of MSP in Specific Body Locations among Medical Imaging Students in the Past 12 Months and the Past 7 Days

MSP in any of the following body region	In the past 12 months n (%)	In the past 7 days n (%)
Neck	23 (36.5)	10 (15.9)
Shoulders	31 (49.2)	8 (12.7)
Elbow	6 (9.5)	0 (0)
Wrist/Hands	13 (20.6)	8 (12.7)
Upper back	17 (27.0)	5 (7.9)
Lower back	24 (38.1)	16 (25.4)
Hips/Thighs	17 (27.0)	3 (4.8)
Knee	14 (22.2)	9 (14.3)
Ankles/Feet	20 (31.7)	3 (4.8)

'n': number of respondents; %: percentage of respondents; MSP: Musculoskeletal pain

Frequency of MSP by Body Region

Table 2 summarises MSP frequency by body region among medical imaging students reporting pain in the past 12 months. The lower back, shoulders, and neck were the most commonly affected regions, whereas elbows and wrists/hands were least reported. For most body regions, MSP was most frequently rated at level 2 on the frequency scale, indicating relatively infrequent episodes, while lower back symptoms were most often rated at level 3 (50.0%, n = 12), suggesting more recurrent pain in this area. Overall, these findings highlight the lower back, shoulders, and neck as priority regions for prevention, likely related to postural and patient-handling demands during clinical training.

Table 2: Frequency of MSP in the Past 12 Months

Body region	n with MSP	Most common frequency response category *, n (%)
Neck	23	2 (Rarely), 11 (47.8)
Shoulders	31	2(Rarely), 14 (45.2)
Elbow	6	2(Rarely),4 (66.7)
Wrist/Hands	13	2(Rarely), 6 (46.2)
Upper back	17	2–3 (Rarely-Sometimes), 6 each (35.3)
Lower back	24	3 (Sometimes), 12 (50.0)

Body region	n with MSP	Most common frequency response category *, n (%)
Hips/Thighs	17	2(Rarely), 8 (47.1)
Knee	14	2(Rarely), 9 (64.3)
Ankles/Feet	20	2(Rarely), 10 (50.0)

'n': number of respondents; %: percentage of respondents; MSP: Musculoskeletal pain; Responses were originally collected on a 5-point Likert scale and recoded into "1-never", "2-rarely", "3-sometimes", "4-often", and "5-always"; Only the most common response category for each body region is shown

Table 3 shows the distribution of MSP in the past 7 days by body region. The lower back had the highest number of students reporting recent pain (n = 16), followed by the neck and knees. When focusing on frequent symptoms (rated often/always), the upper back (60.0%), hips/thighs (66.7%), and lower back (37.5%) showed the greatest proportions, indicating more persistent recent pain in these areas. In contrast, elbow pain was not reported in the past 7 days duration, and neck, knee, and ankle/foot pain were generally less frequent. Overall, recent MSP remained most prominent in the back regions and lower limbs, consistent with the postural and handling demands of clinical training.

Table 3: Frequency of MSP in the Past 7 Days

Body region	n with MSP	Most common frequency response category*, n (%)
Neck	10	2 (Rarely), 3 (30.0)
Shoulders	8	3 (Sometimes),4 (50.0)
Elbow	0	-
Wrist/Hands	8	3-4(Sometimes-Often), 3 each (37.5)
Upper back	5	4 (Often), 3 (60.0)
Lower back	16	4(Often), 6 (37.5)
Hips/Thighs	3	4(Often), 2 (66.7)
Knee	9	2 (Rarely), 6 (66.7)
Ankles/Feet	3	2 (Rarely), 2 (66.7)

'n': number of respondents; %: percentage of respondents; MSP: Musculoskeletal pain; Responses were originally collected on a 5-point Likert scale and recoded into "1-never", "2-rarely", "3-sometimes", "4-often", and "5-always"; Only the most common response category for each body region is shown; – No MSP reported in this body region

Severity of MSP by Body Region

This section reports the severity of MSP, which indicates how intense the pain was in each body region during the specified time frames.

Table 4 describes the most common severity levels of MSP in the past 12 months for each body region among students who reported pain. Overall, symptoms were predominantly mild (category 2) in most regions, including the neck (69.6%), wrists/hands (69.2%), upper back (52.9%), lower back (50.0%), knees (35.7%), and ankles/feet (50.0%). Shoulder pain stood out as generally moderate (category 3), reported by more than half of those with shoulder MSP (54.8%). For hips/thighs, mild and moderate pain were equally common (29.4% each), while elbow symptoms were infrequent and most often rated as no pain (category 1) over the 12-month period. These patterns indicate that, although MSP is widespread across multiple body regions, it is usually of mild to moderate severity rather than severe or very severe.

Table 4: Severity of MSP in the Past 12 Months

Body region	n with MSP	Most common severity response category*, n (%)
Neck	23	2 (Mild), 16 (69.6)
Shoulders	31	3 (Moderate), 17 (54.8)
Elbow	6	1 (No Pain), 3 (50.0)
Wrist/Hands	13	2(Mild), 9 (69.2)
Upper back	17	2(Mild), 9 (52.9)
Lower back	24	2(Mild), 12 (50.0)
Hips/Thighs	17	2-3(Mild-Moderate), 5 each (29.4)
Knee	14	2(Mild), 5 (35.7)
Ankles/Feet	20	2(Mild), 10 (50.0)

'n': number of respondents; %: percentage of respondents; MSP: Musculoskeletal pain; Responses were originally collected on a 5-point Likert scale and recoded into "1-no pain", "2-, mild pain", "3-moderate pain", "4-severe pain", and "5-very severe pain"; Only the most common response category for each body region is shown

Table 5 presents the most common severity levels of MSP in the past 7 days among students who reported recent pain. For several regions, recent symptoms were mainly mild (category 2), including the neck (40.0%), shoulders (37.5%), knees (66.7%), and ankles/feet (66.7%). In contrast, moderate pain (category 3) was most common in the wrists/hands (50.0%), while the upper back showed predominantly severe pain (category 4; 60.0%). Lower back symptoms were split evenly between moderate and severe pain (37.5% each), suggesting more intense recent discomfort in this region. Hips/thighs showed a mixed pattern with one case each of no, severe, and very severe pain. Overall, these findings indicate that, although many students experienced only mild MSP in the previous week, the upper and lower back and wrists/hands were notable sites of more severe recent pain.

Table 5: Severity of MSP in the Past 7 Days

Body region	n with MSP	Most common severity response category, n (%)
Neck	10	2(Mild), 4 (40.0)
Shoulders	8	2(Mild), 3 (37.5)
Elbow	0	-
Wrist/Hands	8	3(Moderate), 4 (50.0)
Upper back	5	4 (Severe), 3 (60.0)
Lower back	16	3-4(Moderate-Severe), 6 each (37.5)
Hips/Thighs	3	1,4,5 (No pain, Severe, Very Severe), 1 each (33.3)
Knee	9	2(Mild), 6 (66.7)
Ankles/Feet	3	2(Mild), 2 (66.7)

'n': number of respondents; %: percentage of respondents; MSP: Musculoskeletal pain; Responses were originally collected on a 5-point Likert scale and recoded into "1-no pain", "2-, mild pain", "3-moderate pain", "4-severe pain", and "5-very severe pain"; Only the most common response category for each body region is shown; – No MSP reported in this body region

Correlation Study

This section addresses the specific research objective to correlate the frequency of MSP in the past 12 months with the potential risk factors associated with MSP in medical imaging students.

Potential Risk Factor Associated With MSP

Table 6 shows that most potential MSP-related risk factors were reported as infrequent, with only a subset of students experiencing them often. The most frequently reported exposure was lower back discomfort during bending, twisting, or lifting patients (31.7%), followed by neck and shoulder discomfort, upper and ankle/foot discomfort, taking short breaks, and practising stress-management techniques (around 20–25%). Overall, these results indicate that postural and handling strains are present during clinical training, but only a minority of students experience them frequently or on a regular basis.

Table 6: Frequency of potential MSP-related risk factors among Medical Imaging students

No.	Risk factor item (summary)	Frequent n (%)	Mean (SD)
1	Neck discomfort when looking up/down for prolonged periods	15 (23.8)	2.73 (1.10)
2	Shoulder discomfort during repetitive tasks	15 (23.8)	2.57 (1.16)
3	Elbow discomfort when lifting/supporting patients	0 (0.0)	1.60 (0.75)
4	Wrist/hand discomfort with forceful or repetitive hand movements	7 (11.1)	2.29 (1.13)
5	Upper back discomfort after maintaining bent-over posture	14 (22.2)	2.37 (1.20)
6	Lower back discomfort when bending, twisting, or lifting patients	20 (31.7)	2.98 (1.09)
7	Hip/thigh discomfort after prolonged standing	4 (6.4)	2.02 (1.06)
8	Knee discomfort when bending or squatting	7 (11.1)	2.17 (1.17)
9	Ankle/foot discomfort after walking or standing for long hours	14 (22.2)	2.38 (1.21)
10	Taking short breaks during extended periods of work	16 (25.4)	2.97 (0.98)
11	Engaging in physical activity/exercise	14 (22.2)	2.81 (1.00)
12	Getting enough sleep (≥7 hours/night)	12 (19.0)	2.70 (1.13)
13	Practising stress-management techniques	15 (23.8)	2.75 (1.15)

Responses were originally collected on a 5-point Likert scale and recoded into "not frequent," "sometimes," and "frequent"; only the "frequent" category is presented here for conciseness

Correlation Analysis

The Kolmogorov–Smirnov test showed that MSP frequency in the past 12 months (Table 2) and all potential risk factors were non-normally distributed (all $p < 0.05$). Consequently, Spearman's correlation was used to examine the relationships between MSP frequency and these risk factors.

Table 7 shows that most significant associations were positive and strong between MSP frequency and discomfort-related ergonomic factors in the corresponding body regions. Strong correlations were observed for neck pain with looking up/down for prolonged periods ($\rho = 0.712$), shoulder pain with repetitive tasks such as equipment handling and patient positioning ($\rho = 0.787$), upper back pain with maintaining a bent-over posture ($\rho = 0.727$), lower back pain with bending, twisting, or lifting patients ($\rho = 0.744$), knee pain with bending or squatting ($\rho = 0.713$), and hips/thigh and ankles/feet pain with prolonged standing or walking during clinical training ($\rho = 0.673$ and $\rho = 0.683$, respectively). Wrist/hand pain also showed a strong correlation with forceful or repetitive hand movements ($\rho = 0.654$), while hips/thigh pain demonstrated a weak but

significant correlation with short breaks during extended work periods (ρ = 0.255). Overall, these findings indicate that more frequent MSP in each body region is closely linked to higher levels of corresponding postural and patient-handling discomfort, underscoring the role of ergonomic demands during clinical training in the development of MSP among medical imaging students.

Table 7: Significant correlations between MSP frequency (past 12 months) and potential risk factors (Spearman's ρ)

Body region	Risk factor item (summary)	Spearman's ρ	Strength*	p-value
Neck	Discomfort due to looking up or down for prolonged periods	0.712	Strong	0.000
Shoulders	Discomfort due to repetitive tasks (e.g., equipment handling, patient positioning)	0.787	Strong	0.000
Elbows	Discomfort when lifting or supporting patients during clinical training	0.503	Moderate	0.000
Wrists/ Hands	Discomfort due to forceful gripping, twisting, or repetitive hand movements	0.654	Strong	0.000
Upper back	Discomfort due to maintaining a bent-over posture during clinical tasks	0.727	Strong	0.000
Lower back	Discomfort due to bending, twisting, or lifting patients	0.744	Strong	0.000
Hips/ Thighs	Discomfort due to prolonged standing during clinical training	0.673	Strong	0.000
Hips/ Thighs	Short break during extended period of work	0.255	Weak	0.044
Knees	Discomfort due to bending or squatting	0.713	Strong	0.000
Ankles/ Feet	Discomfort due to walking or standing for long hours during clinical training	0.683	Strong	0.000

Note: weak 0.20–0.39, moderate 0.40–0.59, strong ≥0.60. Only statistically significant correlations (p ≤ 0.05) are reported; full correlation outputs are available from the corresponding author on reasonable request

DISCUSSION

The study shows that MSP is a prominent problem among medical imaging students, despite most respondents being non-smokers, having a normal BMI, and reporting no previous MSP diagnosis. The participants was predominantly female and drawn from multiple years of study, reflecting varying degrees of clinical exposure. Although only a small proportion reported using ergonomic equipment, MSP was common across several body regions, suggesting that current preventive practices may be insufficient for this group.

Across the body regions examined, the lower back, shoulders, neck, and upper back emerged as the main areas of concern. Using the 12-month recall capturing longer-term or recurrent symptoms, lower back pain showed the highest frequency ratings, with a substantial proportion of students reporting symptoms at intermediate levels on the frequency scale (e.g. "sometimes" or "often"), while shoulders and neck followed closely. In contrast, elbow and wrist/hand symptoms were relatively infrequent and generally rated at lower frequency levels, indicating that these joints may be under comparatively less strain in typical imaging activities.

For the past 7 days duration, lower back pain again featured most prominently, with smaller numbers reporting recent symptoms in the neck, knees, and upper back, highlighting that MSP is not only a historical problem but also a current issue for many students. These findings are broadly consistent with previous work among medical imaging students in other settings, where the lower back, neck, and upper back are also commonly affected. For example, Lorusso et al., (2010) reported that low back pain was the most frequent complaint among Italian medical imaging students, with neck and shoulder symptoms also prevalent. Similarly, a study of undergraduate radiography students in South Africa found a very high prevalence of musculoskeletal disorders (92.4%), with the lower back (79.7%), neck (72.2%), and upper back (54.1%) most frequently affected (Masondo & Khoza, 2023).

In terms of severity, MSP over the past 12 months was predominantly mild across most body regions, particularly in the neck, wrists/hands, upper back, lower back, knees, and ankles/feet, while shoulder symptoms were more often rated as moderate. This pattern suggests that, although MSP is widespread, it is usually not severe over the longer term. However, the 7-day data indicate that recent symptoms can be more intense in specific regions: upper back pain was most commonly severe, and lower back pain was split between moderate and severe levels,

whereas neck, knee, and ankle/foot pain in the past week was mainly mild.

Taken together, the severity and frequency findings indicate a mainly mild-to-moderate, yet recurrent, MSP burden that is strongly concentrated in the spine, especially the lower back, with notable episodes of more severe pain in both the upper and lower back. This recurring pattern suggests that spinal and shoulder regions are consistently vulnerable during radiography training and early clinical exposure. Consequently, these areas should be prioritised in ergonomic prevention strategies, including posture training, safer patient-handling techniques, and adjustments to clinical work routines.

The analysis of potential risk factors in this study highlights task-related and postural exposures as key contributors to MSP. Strong positive correlations were observed between neck pain and activities involving prolonged neck flexion/extension, shoulder pain and repetitive shoulder tasks, upper and lower back pain and bending, twisting, or lifting patients, and lower-limb pain and prolonged standing or walking during clinical work. These moderate to strong correlations supports previous evidence that physical workload and awkward postures are central drivers of MSP among radiography students and practitioners.

Additional correlation analyses (not shown in tables) indicated that lifestyle-related factors such as physical activity, sleep duration, and stress-management practices did not show significant correlations with MSP frequency in this cohort, suggesting that clinical task demands and work posture may exert a stronger influence on MSP than general lifestyle habits during training.

The data for potential risk factors also provide useful context. While most students reported many exposures as infrequent, a substantial subgroup experienced certain high-risk tasks regularly, particularly lower back discomfort during bending, twisting, or lifting patients, and discomfort related to prolonged standing or walking. This aligns with broader literature on healthcare and allied health students, which consistently links manual handling, static standing, and constrained postures to MSP and work-related musculoskeletal disorders.

CONCLUSION

The findings of this study have several practical implications. First, they underscore the importance of integrating structured ergonomics education into medical imaging curricula, with explicit emphasis on safe patient handling, posture optimisation, and micro-break strategies during imaging procedures. Second, clinical training sites should consider workload management and simple environmental adjustments (e.g. equipment positioning, use of aids for patient transfer) to reduce static loading and awkward postures. Third, early identification and support for students reporting frequent or severe MSP, particularly in the lower back and upper back, may help prevent progression to chronic symptoms as they enter the workforce.

This study has some limitations that must be acknowledged. Data were collected over a relatively short period and from a single institution, resulting in a modest sample size that may limit the precision and generalisability of estimates. The cross-sectional, self-report design is also susceptible to recall bias and individual differences in pain perception, and causal relationships between MSP and the identified risk factors cannot be inferred. Nevertheless, the use of a validated MSP instrument with 12-month and 7-day recall, alongside a structured set of task-related and postural exposure items, strengthens the internal validity of the findings and provides a more detailed picture of MSP patterns than many previous studies.

Future research should build on these results by employing longitudinal designs to track MSP trails across academic years and clinical rotations, and by evaluating the impact of targeted ergonomic and workload interventions over time. Experimental or quasi-experimental studies testing posture and handling training, ergonomic modifications, and structured physical-activity programmes would help determine which strategies are most effective in reducing MSP among medical imaging students. Larger, multi-institution cohorts would further enhance generalisability and allow comparison across different training environments and clinical settings.

ACKNOWLEDGEMENT

This research received no fund from any grant. Artificial intelligence tools were used to assist with content development and language editing in the preparation of this manuscript. The authors retain full responsibility for the final content and affirm its integrity and accuracy.

REFERENCES

Abdulaziz, A. A., Althaqafi, A. M., Hindi, A. M., Khan, S. A., Atalla, A. A., & Hendi, O. M. (2019). Prevalence of musculoskeletal disorders and its correlation to physical activity among health specialty students. *International Journal of Preventive Medicine*, *10*(1), 19–24. https://doi.org/10.4103/IJPVM.IJPVM_436_18

Alias, A. N., Karuppiah, K., How, V., & Perumal, V. (2020). Does Prolonged Standing at Work Among Teachers Associated with Musculoskeletal Disorders (MSDs)? In *Malaysian Journal of Medicine and Health Sciences* (Vol. 16, Issue 2).

Bonett, D. G., & Wright, T. A. (2015). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, *36*(1), 3–15. https://doi.org/10.1002/JOB.1960

Dawson, A. P., Steele, E. J., Hodges, P. W., & Stewart, S. (2009). Development and Test-Retest Reliability of an Extended Version of the Nordic Musculoskeletal Questionnaire (NMQ-E): A Screening Instrument for Musculoskeletal Pain. *Journal of Pain*, *10*(5), 517–526. https://doi.org/10.1016/j.jpain.2008.11.008

El-Tallawy, S. N., Nalamasu, R., Salem, G. I., LeQuang, J. A. K., Pergolizzi, J. V., & Christo, P. J. (2021). Management of Musculoskeletal Pain: An Update with Emphasis on Chronic Musculoskeletal Pain. *Pain and Therapy*, *10*(1), 181–209. https://doi.org/10.1007/S40122-021-00235-2/METRICS

Felemban, R. A., Sofi, R. A., Alhebshi, S. A., Alharbi, S. G., Farsi, N. J., Abduljabbar, F. H., & Farsi, J. M. A. (2021). Prevalence and predictors of musculoskeletal pain among undergraduate students at a dental school in Saudi Arabia. *Clinical, Cosmetic and Investigational Dentistry*, *13*, 39–46. https://doi.org/10.2147/CCIDE.S292970

Kuorinka, I., Jonsson, B., Kilbom, A., Vinterberg, H., Biering-Sørensen, F., Andersson, G., & Jørgensen, K. (1987). Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. *Applied Ergonomics*, *18*(3), 233–237. https://doi.org/10.1016/0003-6870(87)90010-X

Laerd Statistics. (2018). *Pearson's product-moment correlation using SPSS Statistics.* Statistical Tutorials and Software Guides. Retrieved from https://statistics.laerd.com/

Lee, J., & Cho, J.-H. (2012). Survey of the musculoskeletal disorders of radiological technologists. *Journal of the Korean Society of Radiology*, *6*(1), 53–61. https://doi.org/10.7742/JKSR.2012.6.1.053

Legault, É. P., Cantin, V., & Descarreaux, M. (2014). Assessment of musculoskeletal symptoms and their impacts in the adolescent population: Adaptation and validation of a questionnaire. *BMC Pediatrics*, *14*(1). https://doi.org/10.1186/1471-2431-14-173

Lorusso, A., Vimercati, L., & L'Abbate, N. (2010). Musculoskeletal complaints among Italian x-ray technology students: A cross-sectional questionnaire survey. *BMC Research Notes*, *3*(1), 114-. https://doi.org/10.1186/1756-0500-3-114/TABLES/3

Masondo, S. S., & Khoza, T. E. (2023). Musculoskeletal disorders among undergraduate radiography students during work integrated learning in Kwazulu-Natal, South Africa. *Journal of Medical Imaging and Radiation Sciences*, *54*(1), 153–161. https://doi.org/10.1016/J.JMIR.2022.12.008

Rasmussen, C. D. N., Holtermann, A., & Jørgensen, M. B. (2018). Recall Bias in Low Back Pain among Workers. *Spine*, *43*(12), E727–E733. https://doi.org/10.1097/BRS.0000000000002457

Senarath, M., Thalwaththe, S., & Tennakoon, S. (2021). Prevalence of Selected Musculoskeletal Disorders among the Students of Faculty of Allied Health Sciences, University of Peradeniya. *Journal of Musculoskeletal Disorders and Treatment*, *7*(2), 097. https://doi.org/10.23937/2572-3243.1510097

Syafiqah, N., Kamalruzaman, A., Adilah, T., Sabri, T., Nor, S., & Isa, I. (2021). Musculoskeletal disorders and quality of life among undergraduate health sciences students: A cross-sectional study. *Healthscope: The Official Research Book of Faculty of Health Sciences, UiTM*, *4*(1), 99–105. https://healthscopefsk.com/index.php/research/article/view/231

Umar, L., Abdullahi, U., Kaka, B., & Aliyu, F. H. (2025). Prevalence and multifactorial determinants of musculoskeletal pain among family caregivers of stroke survivors in Northwestern Nigeria: a cross-sectional analysis with implications for policy and practice. *Discover Public Health*, *22*(1). https://doi.org/10.1186/s12982-025-01025-3

ORIGINAL ARTICLE

Translation and Cultural Adaptation of the Malay Version of the Noise Exposure Questionnaire (M-NEQ)

Noraidah Ismail[1*], Nur Alia Natasha Binti Apandi[1]

[1]*Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia*

Keywords:
Noise-induced hearing loss (NIHL); Noise Exposure Questionnaire (NEQ); Malay language adaptation; Cross-cultural translation; Questionnaire validation; Content validity; Face validity

ABSTRACT

Background: Noise-induced hearing loss (NIHL) remains a significant global health issue, yet self-assessment tools for individual noise exposure are limited in Malaysia. The Noise Exposure Questionnaire (NEQ), originally developed by Johnson et al., has been validated in multiple languages but has not been adapted for Malay speakers. This study aimed to translate, culturally adapt, and validate the Malay version of the NEQ (M-NEQ) for use among the Malaysian population.
Methods: A cross-sectional study was conducted involving forward–backward translation, expert review, and validation phases. Two independent forward and backward translations were harmonized through consensus meetings. Content validation was performed by eight expert audiologists using a four-point relevance scale, while face validation involved ten lay participants from diverse backgrounds assessing clarity and comprehensibility. Item-level and scale-level Content Validity Index (I-CVI and S-CVI/Ave) and Face Validation Index (I-FVI and S-FVI/Ave) were calculated.
Results: The I-CVI values ranged from 0.87 to 1.00, with an overall S-CVI/Ave of 0.98, indicating excellent content validity. Face validation yielded I-FVI scores between 0.90 and 1.00, and an S-FVI/Ave of 0.97, demonstrating high clarity and comprehensibility. Minor revisions were made to enhance linguistic precision and cultural relevance.
Conclusion: The Malay version of the NEQ (M-NEQ) was successfully adapted and validated, showing strong content and face validity. It represents a promising tool for assessing daily noise exposure in occupational and recreational settings among Malaysian adults. Further research is recommended to evaluate its reliability and construct validity in broader populations.

INTRODUCTION

Noise-induced hearing loss (NIHL) is a growing global health concern, resulting from both occupational and recreational sources of noise exposure. According to the World Health Organization (WHO, 2024), over 1.5 billion people worldwide are affected by hearing loss, with projections indicating that more than 700 million individuals may experience disabling hearing loss by 2050. NIHL alone impacts approximately 5% of the global population. To mitigate these risks, regulatory bodies such as the Occupational Safety and Health Administration (OSHA) 2019 recommend limiting daily noise exposure to an average of 85 decibels over an eight-hour workday.

While industrial environments are commonly associated with hearing loss, increasing evidence highlights the risks posed by everyday activities such as listening to music at high volumes through headphones, attending concerts, and exposure to urban noise from public transportation (Neitzel & Fligor, 2019; Osmanoğlu et al., 2024). These non-occupational sources contribute significantly to cumulative noise exposure, underscoring the need for accurate individual-level assessment tools to identify NIHL risk and support preventive strategies.

However, tracking personal noise exposure over time remains challenging due to the limited availability of reliable self-assessment instruments. In response to this gap, Johnson et al. (2017) developed the Noise Exposure Questionnaire (NEQ), a standardized tool designed to estimate annual noise exposure (ANE) from both occupational and recreational sources. The NEQ has been successfully adapted into several languages, including Mandarin (Han et al., 2022) and Portuguese (Oliveira et al., 2023), with studies consistently demonstrating its strong validity and reliability.

Given the effectiveness of the NEQ in diverse populations, this study aims to translate, culturally adapt, and validate a Malay-language version of the NEQ (M-NEQ) for use among Malay speakers in Malaysia. The adaptation seeks

* *Corresponding author.*
E-mail address: noraidah@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

to ensure linguistic clarity, cultural relevance, and applicability to everyday noise experiences in both occupational and non-occupational contexts.

To date, no self-reported tools for evaluating daily noise exposure have been adapted specifically for the Malay-speaking population. Although Razali et al. (2022) developed and validated a combined noise and chemical exposure questionnaire for hospital workers in Malaysia, their instrument is tailored to occupational settings and includes chemical exposure components. In contrast, the present study focuses exclusively on noise exposure and is designed for broader use across the general population, addressing a critical gap in assessing environmental and lifestyle-related noise exposure in the Malaysian context.

MATERIALS AND METHODS

Study Design

This study employed a cross-sectional design, allowing data collection at a single point in time from diverse participant groups. This approach was appropriate for evaluating the quality, clarity, and relevance of the translated questionnaire without manipulating variables or tracking participants longitudinally (Wang & Cheng, 2020). The primary objective was to ensure that the Malay version of the Noise Exposure Questionnaire (M-NEQ) was linguistically accurate, culturally appropriate, and comprehensible to the target population.

Study Population and Sample Size

The translation of the M-NEQ followed a structured process involving two forward translators and two backward translators. The forward translators, both with audiology backgrounds, ensured accurate translation of technical terminology. In contrast, the backward translator, without an audiology background, provided an unbiased retranslation into the source language (Tsang et al., 2017). All translators were proficient in both Malay and English.

To ensure consistency, two audiology experts with experience in noise-induced hearing loss (NIHL) harmonized the forward translations. A harmonization meeting was conducted with the researcher, both audiology experts, and one representative from each translation team to review the translated items, resolve discrepancies, and ensure fidelity to the original questionnaire.

Content validation was conducted by eight audiologists, each with over five years of professional experience, including three with direct experience in NIHL. For face validation, ten participants from the target population were recruited, including dentists, occupational safety and health (OSH) personnel, engineering academics, and university students who reported daily use of personal listening devices at approximately 50% of maximum volume. All participants were proficient in both Malay and English.

Sampling Method

A purposive sampling strategy was employed to recruit individuals who met specific criteria relevant to the study. This method ensured the inclusion of participants with adequate language proficiency and domain knowledge for the translation and validation processes. It also facilitated the collection of meaningful and relevant feedback (Palinkas et al., 2015).

Instrumentation

This study involved the translation and validation of the Noise Exposure Questionnaire (NEQ), originally developed by Johnson et al. (2017) in English. The NEQ comprises 11 items designed to estimate individual exposure to hazardous noise from occupational and recreational sources. Items assess typical noise exposure behaviors, such as the use of power tools, attendance at loud events, and use of personal listening devices—factors contributing to NIHL.

Study Procedure

This study consisted of three main phases: (i) adaptation, (ii) translation, and (iii) validation, as illustrated in Figure 1.

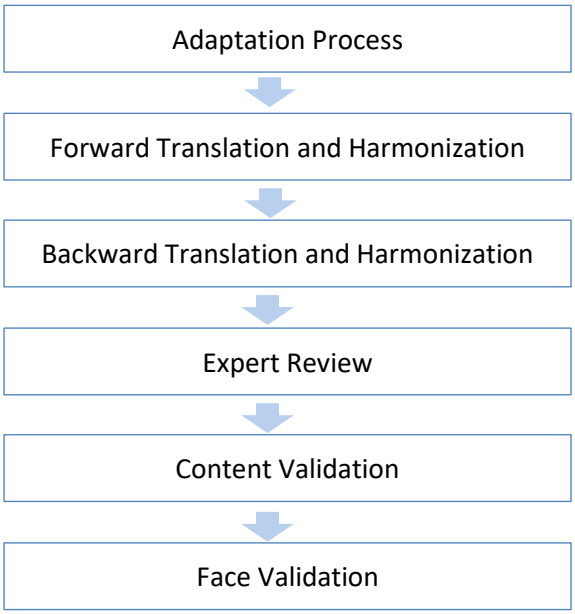


Figure 1: The study procedure flowchart

Adaptation Process

At the outset of the study, an adaptation process was undertaken to modify the original questionnaire to ensure cultural relevance and appropriateness for the Malaysian context. It involved omitting, rephrasing, and combining certain questionnaire items. The original Noise Exposure Questionnaire (NEQ) was systematically reviewed in consultation with the research supervisor. All modifications were mutually agreed upon during the review process. The final adapted version was compiled using Microsoft Excel in preparation for the subsequent translation phase.

Translation Process

The translation process was conducted in five stages: (i) forward translation, (ii) harmonization of the forward translations, (iii) backward translation, (iv) harmonization of the backward translations, and (v) expert review of both forward and backward translations to resolve discrepancies and ensure conceptual equivalence with the original version.

Forward Translation and Harmonization

The original Noise Exposure Questionnaire was independently translated into the Malay language by two forward translators. The NEQ was compiled into an Excel document and was distributed to the professional translators via email. Each translator submitted their respective versions via the researcher's email for review. To harmonize the translations, the researcher compared both versions and identified differences in word choices and phrasing. The most contextually appropriate terms were selected to ensure clarity and alignment with the original meaning. A series of discussions was subsequently held with the research supervisor to review the selected terms and resolve any uncertainties. Amendments were made based on mutual agreement. As a result, a harmonized Malay version of the forward translation was finalized for the next phase of the study.

Backward Translation and Harmonization

During this stage, the harmonized forward-translated questionnaire was independently translated back into English by two professional backward translators who were not from an audiology background and were unfamiliar with the original questionnaire. Similarly, the harmonized Malay questionnaire was compiled into an Excel document and emailed to the professional backward translators. Both translators submitted their translations via the researcher's email. The researcher compared both backward translations and identified variations in wording. The most contextually accurate and semantically equivalent terms were selected to ensure the meaning closely reflected that of the original English version. A series of discussions was subsequently held with the research supervisor to verify that the chosen wording accurately represented the intent of the original questionnaire. Amendments were made based on mutual agreement, and the harmonized backward-translated version was finalized for expert review.

Expert Committee Review

A final comparison between the original English version, harmonized forward-translated, harmonized backward-translated questionnaire was conducted during an online meeting with the expert panels. The purpose of this review was to identify and resolve any discrepancies between the original and translated versions to ensure conceptual and linguistic equivalence. Throughout this process, the expert panels discussed differences in wording and interpretation. Revisions, including the rephrasing of certain items and the addition of clarifying terms, were made where necessary to enhance accuracy, clarity, and cultural appropriateness. The review concluded when a consensus was reached on all items, resulting in a finalized version of the Malay Noise Exposure Questionnaire (M-NEQ) ready for validation.

Validation

Finally, the questionnaire was prepared for content and face validation. The validity of the questionnaire was assessed through content and face validation processes. The Malay translated questionnaire was compiled in

Microsoft Word format and distributed to the validators via email. Clear instructions in Malay were provided to facilitate understanding. The returned documents were then analyzed based on the validators' scores and qualitative feedback.

Content Validation

A total of eight content validators were chosen based on the proposed inclusion criteria. Their task was to rate the relevance of the questionnaire items using the following scoring system; i) 1-irrelevant, ii) 2-somewhat relevant, iii) 3-quite relevant, and iv) 4-highly relevant. This is done to test the relevance of the components and to ensure that the items accurately reflect the domain. After the content validation survey form was received back from the content validators, the Content Validity Index (CVI) at item level (item-level content validity index (I-CVI)) and at scale level (scale-level content validity index (S-CVI)) were calculated. In addition, content validators' feedback was also considered to improve the items in the questionnaire further.

Face Validation

A total of ten target users were recruited as face validators based on the proposed inclusion criteria. Their task was to rate the clarity and comprehensibility of the questionnaire components using the following scoring system: i) 1-not clear and cannot be understood, ii) 2-somewhat clear and can be understood, iii) 3-clear and easily understood, and iv) 4-very clear and very easy to understand. After the face validators returned the face validation survey forms, the Face Validation Index (FVI) at the item level (item-level face validation index (I-FVI)) and the scale level (scale-level face validity index (S-FVI)) were calculated. Modifications to the components were also conducted based on suggestions from the face validators.

Data Analysis

Content Validity Index (CVI)

The I-CVI and S-CVI were calculated according to the score given by the experts. Questionnaire components with a score of 3 and 4 were considered relevant and equal to 1, whilst a score of 1 or 2 were considered irrelevant and equal to 0. According to Yusoff (2019), for a panel of at least six experts, an I-CVI value of 0.83 or higher is considered satisfactory. Similarly, Polit & Beck (2006) recommend an acceptable I-CVI threshold of not less than 0.83 when six or more experts are involved. Meanwhile, the acceptable value for S-CVI/Ave was recommended of 0.9 or higher.

Equation for I-CVI:

$$\text{I-CVI} = \frac{The\ number\ of\ experts\ giving\ rating\ 3\ or\ 4}{Total\ number\ of\ experts}$$

Equation for S-CVI/Ave:

$$\text{S-CVI} = \frac{The\ sum\ of\ I{-}CVI}{Total\ number\ of\ items}$$

Face Validation Index (FVI)

The I-FVI and S-FVI were calculated according to the score given by the experts. For each questionnaire item, scores of 3 and 4 were considered relevant and assigned a value of 1, while scores of 1 or 2 were considered irrelevant and assigned a value of 0. An I-FVI greater than 0.80 and an S-FVI exceeding 0.83 would be considered as acceptable levels of face validity (Rahman et al. (2021).

Equation for I-FVI

$$\text{I-FVI} = \frac{The\ number\ of\ experts\ giving\ rating\ 3\ or\ 4}{Total\ number\ of\ experts}$$

Equation for S-FVI/Ave:

$$\text{S-FVI} = \frac{The\ sum\ of\ I{-}FVI}{Total\ number\ of\ items}$$

RESULTS

Adaptation Process

During the adaptation process, several changes were made to the original questionnaire to ensure cultural appropriateness for the Malaysian context. These changes involved omitting, rephrasing, and combining certain item.

Under the omission category, the word "snowmobiles" in Item 4a was removed as it is not relevant or commonly used in Malaysia. No replacement was made since the other listed examples such as motorcycles, jet skis, and speed boats were already sufficient to convey the intended meaning.

Items 10 and 11, which originally referred to noisy paid jobs during and outside the summer months, were rephrased and combined into a single item. This adaptation was necessary due to the absence of a summer break concept in the Malaysian context. The revised item refers to any noisy paid work conducted within the past year, including work undertaken during school holidays or other free time.

These changes were made while maintaining the original intent of the questionnaire, ensuring that the revised

items were culturally relevant, easier to understand, and applicable to the local context.

Table 1: Adaptation categories

Adaptation categories	Items	Changes	
		From	To
Omission	3a	How often did you ride/operate motorized vehicles such as motorcycles, jet skis, speed boats, snowmobiles, or four-wheelers?	How often did you ride/operate motorized vehicles such as motorcycles, jet skis, speed boats, snowmobiles, or four-wheelers?
	11a	Other than during the summer, over the past year, did you work one or more noisy paid jobs, such as in construction, farming, a factory, lawn service, carwash, or other indoor or outdoor job working around loud equipment or machinery? By noisy job, we mean sounds so loud that you had to shout or speak in a raised voice to be heard at arm's length.	[Merged with Q10] Adapted into one general question covering "the past year", without seasonal distinction.
Rephrasing	10a	Now think back to this past summer. Over the summer months, did you work a noisy paid job, such as in construction, farming, a factory, lawn service, carwash, or other indoor or outdoor job working around loud equipment or machinery? By noisy job, we mean sounds so loud that you had to shout or speak in a raised voice to be heard at arm's length.	In the past year, did you work a noisy paid job, such as in construction, farming, a factory, lawn service, carwash, or other indoor or outdoor job working around loud equipment or machinery? By noisy job, we mean sounds so loud that you had to shout or speak in a raised voice to be heard at arm's length.

Translation and Harmonization Process

The translation process successfully resulted in a harmonized and culturally adapted Malay version of the NEQ. During the expert review, several items were identified for improvement. The expert panel recommended revisions mainly involving rewording for semantic clarity and alignment with Malaysian linguistic norms. As a result, changes were made to multiple items across the questionnaire. These revisions helped improve the comprehensibility of the items for the public, especially in the context of self-reporting noise exposure. The changes made to the following expert review are presented in Tables 2 and 3 below.

Table 2: Original NEQ, Harmonized Malay translated, Harmonized English Translated, and Harmonized Malay following expert review

Items	Original NEQ	Harmonized Forward Translation	Harmonized Backward Translation	Harmonized Forward Translation (2)
1a	Outside of a paid job, how often did you use power tools, chainsaws, or other shop tools?	Di luar waktu bekerja, berapa kerapkah anda menggunakan peralatan kuasa elektrik seperti gergaji elektrik atau alat bengkel lain?	Outside of a paid job, how often did you use power tools such as electric saws or other workshop equipment?	Di luar waktu bekerja, berapa kerapkah anda menggunakan peralatan kuasa elektrik seperti gergaji elektrik atau alat bengkel lain?
1b	If you used power tools, on average, how many hours did each time/session last?	Jika anda menggunakan peralatan tersebut, berapa lamakah sesi penggunaannya?	If you use power tools, how long did each usage last?	Jika anda menggunakan peralatan kuasa elektrik tersebut, berapa lamakah sesi penggunaannya?

1c	If you used power tools, how often did you wear earplugs or earmuffs during this activity?	Jika anda menggunakan peralatan kuasa elektrik, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?	If you use power tools, how often did you wear hearing protection during this activity?	Jika anda menggunakan peralatan kuasa elektrik, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?
2a	Outside of a paid job, how often did you drive heavy equipment or use loud machinery (such as tractors, trucks, or farming or lawn equipment like mowers/leaf blowers)?	Di luar waktu bekerja, berapa kerapkah anda memandu atau menggunakan jentera berat seperti traktor, trak, atau peralatan pertanian/penjagaan rumput seperti mesin pemotong rumput/pam daun?	Outside of a paid job, how often did you drive or operate heavy machinery such as tractors, trucks, or agricultural/landscaping equipment like lawnmowers?	Di luar waktu bekerja, berapa kerapkah anda memandu atau menggunakan jentera berat seperti traktor, trak, atau peralatan perladangan/penjagaan rumput seperti mesin pemotong rumput/pam daun?
2b	If you drove/used loud machinery, on average, how many hours did each time/session last?	Jika anda memandu atau menggunakan jentera berat tersebut, berapa lamakah sesi penggunaannya?	If you drove or operated heavy machinery, how long did each usage last?	Jika anda memandu atau menggunakan jentera berat tersebut, berapa lamakah sesi penggunaannya?
2c	If you drove/used machinery, how often did you wear earplugs or earmuffs during this activity?	Jika anda memandu atau menggunakan jentera berat, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?	If you drove or operate heavy machinery, how often do you wear hearing protection during this activity?	Jika anda memandu atau menggunakan jentera berat, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?
3a	How often did you attend car/truck races, commercial/high school sporting events, music concerts/dances or any other events with amplified public announcement (PA)/music systems?	Berapa kerapkah anda menghadiri acara seperti perlumbaan kereta/lori, acara sukan, konsert muzik, atau acara lain dengan sistem pembesar suara?	How often did you attend events such as car/truck races, sports events, music concerts, or other events with amplified public announcement (PA)/music systems?	Berapa kerapkah anda menghadiri acara seperti perlumbaan kereta/lori, acara sukan, konsert muzik, atau acara lain yang menggunakan sistem pembesar suara?
3b	If you attended these events, on average, how many hours did each time/session last?	Jika anda menghadiri acara ini, berapa lamakah sesi tersebut berlangsung?	If you attended these events, how long did they usually last?	Jika anda menghadiri acara ini, berapa lamakah sesi tersebut berlangsung?
3c	If you attended these events, how often did you wear earplugs or earmuffs during this activity?	Jika anda menghadiri acara ini, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?	If you attended these events, how often did you wear hearing protection during these activities?	Jika anda menghadiri acara ini, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?
4a	How often did you ride/operate motorized vehicles such as motorcycles, jet skis, speed boats, snowmobiles, or four-wheelers?	Berapa kerapkah anda menggunakan kenderaan bermotor seperti motosikal, jet ski, bot laju, atau motor bertayar empat?	How often did you use motor vehicles such as motorcycles, jet skis, speedboats, or all-terrain vehicles?	Berapa kerapkah anda menggunakan kenderaan bermotor seperti motosikal, jet ski, bot laju, atau kenderaan empat roda?
4b	If you rode motorized vehicles, on average, how many hours did each time/session last?	Jika anda menggunakan kenderaan bermotor tersebut, berapa lamakah sesi penggunaannya?	If you use these motor vehicles, how long did each session last?	Jika anda menggunakan kenderaan bermotor tersebut, berapa lamakah sesi penggunaannya?
4c	If you rode motorized vehicles, how often did you wear earplugs or earmuffs during this	Jika anda menggunakan kenderaan bermotor, berapa kerapkah anda memakai alat pelindung	If you use motor vehicles, how often do you wear hearing protection during this activity?	Jika anda menggunakan kenderaan bermotor, berapa kerapkah anda memakai alat pelindung

	activity?	pendengaran semasa aktiviti ini?		pendengaran semasa aktiviti ini?
5a	How often did you ride in or pilot small aircraft/private airplanes?	Berapa kerapkah anda menaiki atau membawa pesawat kecil/pesawat peribadi?	How often did you ride in or pilot a small aircraft/private plane?	Berapa kerapkah anda menaiki atau membawa pesawat kecil/pesawat peribadi?
5b	If you flew airplanes, on average, how many hours did each time/session last?	Jika anda menaiki pesawat, berapa lamakah sesi penerbangan tersebut?	If you ride in an aircraft, how long did the flight session last?	Jika anda menaiki pesawat, berapa lamakah sesi penerbangan tersebut?
5c	If you flew airplanes, how often did you wear earplugs or earmuffs during this activity?	Jika anda menaiki pesawat, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?	If you ride in an aircraft, how often did you wear hearing protection during this activity?	Jika anda menaiki pesawat, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?
6a	How often were you around or did you shoot firearms such as rifles, pistols, shotguns, etc.?	Berapa kerapkah anda menggunakan atau berada berhampiran senjata api seperti senapang,, pistol, senapang patah atau sebagainya?	How often did you use or be near firearms such as rifles, pistols, shotguns, or similar weapons?	Berapa kerapkah anda menggunakan atau berada berhampiran senjata api seperti senapang, pistol, senapang patah atau sebagainya?
6b	If you were around/shot firearms, on average, how many shots did you fire each time/session?	Jika anda menggunakan senjata api tersebut, berapa banyakkah tembakan yang anda lepaskan semasa sesi itu?	If you used firearms, how many shots did you fire during the session?	Jika anda menggunakan senjata api tersebut, berapa banyakkah tembakan yang anda lepaskan semasa sesi itu?
6c	If you were around/shot firearms, how often did you wear earplugs or earmuffs while shooting?	Jika anda menggunakan senjata api tersebut, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?	If you used these firearms, how often do you wear hearing protection while shooting?	Jika anda menggunakan senjata api tersebut, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?
7a	How often did you play a musical instrument?	Berapa kerapkah anda bermain alat muzik?	How often did you play a musical instrument?	Berapa kerapkah anda bermain alat muzik?
7b	If you played, please tell us what musical instrument:	Jika anda bermain alat muzik, sila nyatakan alat muzik tersebut:	If you played a musical instrument, please specify the instrument:	Jika anda bermain alat muzik, sila nyatakan alat muzik tersebut:
7c	If you played a musical instrument, on average, how many hours did each time/session last?	Jika anda bermain alat muzik, berapa lamakah setiap sesi berlangsung?	If you played a musical instrument, how long did each session last?	Jika anda bermain alat muzik, berapa lamakah setiap sesi berlangsung?
7d	If you played a musical instrument, how often did you wear earplugs or earmuffs while playing?	Jika anda bermain alat muzik, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?	If you played a musical instrument, how often do you wear hearing protection while playing?	Jika anda bermain alat muzik, berapa kerapkah anda memakai alat pelindung pendengaran semasa aktiviti ini?
8a	How often did you listen to music, radio programs, etc. using personal headsets or earphones?	Berapa kerapkah anda mendengar muzik, program radio, dan lain-lain menggunakan “headphone” atau “earphone”?	How often did you listen to music, radio programmes, or other audio using headphones or earphones?	Berapa kerapkah anda mendengar muzik, program radio, dan lain-lain menggunakan fon kepala atau fon telinga?
8b	If you listened through earphones, on average, how many hours did each time/session last?	Jika anda mendengar melalui “headphone” atau “earphone”, secara purata, berapa lamakah setiap sesi berlangsung?	If you listened using headphones or earphones, how long did each session last?	Jika anda mendengar melalui fon kepala atau fon telinga, berapa lamakah setiap sesi berlangsung?
9a	Other than music concerts and headset use (already	Selain konsert muzik dan penggunaan “headphone”	Besides music concerts and headphone use (as	Selain konsert muzik dan penggunaan fon kepala

	covered in questions 3 and 8), how often did you listen to music, radio programs, etc. from audio speakers in a car or at home?	(seperti yang disebut dalam soalan 3 dan 8), berapa kerapkah anda mendengar muzik, program radio, dan lain-lain melalui pembesar suara di rumah atau dalam kereta?	mentioned in questions 3 and 8), how often did you listen to music, radio programmes, and other audio through speakers at home or in the car?	(seperti yang disebut dalam soalan 3 dan 8), berapa kerapkah anda mendengar muzik, program radio, dan lain-lain melalui pembesar suara di rumah atau dalam kereta?
9b	If you listened via speakers, on average, how many hours did each time/session last?	Jika anda mendengar melalui pembesar suara, berapa lamakah setiap sesi berlangsung?	If you listened through speakers, how long did each session last?	Jika anda mendengar melalui pembesar suara, berapa lamakah setiap sesi berlangsung?
10a	Now think back to this past summer. Over the summer months, did you work a noisy paid job, such as in construction, farming, a factory, lawn service, carwash, or other indoor or outdoor job working around loud equipment or machinery? By noisy job, we mean sounds so loud that you had to shout or speak in a raised voice to be heard at arm's length.	Adakah anda bekerja dalam keadaan terdedah kepada bunyi yang bising, seperti di tapak pembinaan, pertanian, kilang, perkhidmatan landskap, mencuci kereta, atau pekerjaan lain yang melibatkan peralatan atau mesin yang kuat? (Dengan pekerjaan bising, kami maksudkan bunyi yang begitu kuat sehingga anda perlu menjerit atau bercakap dengan suara kuat untuk didengari pada jarak yang dekat)	Did you work in an environment exposed to loud noise, such as at a construction site, farm, factory, landscaping service, car wash, or other jobs involving loud equipment or machinery? (By noisy work, we mean sounds so loud that you have to shout or speak loudly to be heard at close range.)	Adakah anda bekerja dalam keadaan terdedah kepada bunyi yang bising, seperti di tapak pembinaan, perladangan, kilang, perkhidmatan mesin rumput, mencuci kereta, atau pekerjaan lain yang melibatkan peralatan atau mesin yang kuat? (Dengan pekerjaan bising, kami maksudkan bunyi yang sangat kuat sehingga anda perlu menjerit atau bercakap dengan suara kuat untuk didengari pada jarak yang dekat)
10b	If yes, please describe this noisy job:	Jika ya, sila nyatakan pekerjaan tersebut:	If yes, please specify the noisy job:	Jika ya, sila nyatakan pekerjaan tersebut:
10c	If you worked a noisy job, please estimate the number of hours you worked in a typical week:	Jika anda bekerja di tempat yang bising, sila anggarkan jumlah jam anda bekerja dalam satu minggu:	If you worked in a noisy environment, please estimate the number of hours you work per week:	Jika anda bekerja di tempat yang bising, sila anggarkan jumlah jam anda bekerja dalam satu minggu:
10d	If you worked a noisy job this summer, did your employer give you earplugs or earmuffs to wear at work?	Adakah majikan anda menyediakan alat pelindung pendengaran untuk digunakan semasa bekerja?	Did your employer provide hearing protection equipment for use while working?	Adakah majikan anda menyediakan alat pelindung pendengaran untuk digunakan semasa bekerja?
10e	How often did you wear earplugs or earmuffs when around loud noise at this summer job?	Berapa kerapkah anda memakai alat pelindung pendengaran semasa bekerja dalam persekitaran yang bising?	How often did you wear hearing protection while working in a noisy environment?	Berapa kerapkah anda memakai alat pelindung pendengaran semasa bekerja dalam persekitaran yang bising?

Table 3: Changes in translation made following the expert reviewal

Items	Harmonized Forward Translation	Harmonized Forward Translation (2)	Changes	
			From	To
2a	Di luar waktu bekerja, berapa kerapkah anda memandu atau menggunakan jentera berat seperti traktor, trak, atau peralatan pertanian/penjagaan rumput seperti mesin pemotong rumput/pam daun?	Di luar waktu bekerja, berapa kerapkah anda memandu atau menggunakan jentera berat seperti traktor, trak, atau peralatan perladangan/penjagaan rumput seperti mesin pemotong rumput/pam daun?	pertanian	perladangan
3a	Berapa kerapkah anda menghadiri acara seperti perlumbaan kereta/lori,	Berapa kerapkah anda menghadiri acara seperti perlumbaan kereta/lori, acara	dengan	yang menggunakan

	acara sukan, konsert muzik, atau acara lain dengan sistem pembesar suara?	sukan, konsert muzik, atau acara lain yang menggunakan sistem pembesar suara?		
4a	Berapa kerapkah anda menggunakan kenderaan bermotor seperti motosikal, jet ski, bot laju, atau motor bertayar empat?	Berapa kerapkah anda menggunakan kenderaan bermotor seperti motosikal, jet ski, bot laju, atau kenderaan empat roda?	motor bertayar empat	kenderaan empat roda
8a	Berapa kerapkah anda mendengar muzik, program radio, dan lain-lain menggunakan "headphone" atau "earphone"?	Berapa kerapkah anda mendengar muzik, program radio, dan lain-lain menggunakan fon kepala atau fon telinga?	'headphone' atau 'earphone'	fon kepala atau fon telinga
9a	Selain konsert muzik dan penggunaan "headphone" (seperti yang disebut dalam soalan 3 dan 8), berapa kerapkah anda mendengar muzik, program radio, dan lain-lain melalui pembesar suara di rumah atau dalam kereta?	Selain konsert muzik dan penggunaan fon kepala (seperti yang disebut dalam soalan 3 dan 8), berapa kerapkah anda mendengar muzik, program radio, dan lain-lain melalui pembesar suara di rumah atau dalam kereta?	'headphone'	fon kepala
10a	Adakah anda bekerja dalam keadaan terdedah kepada bunyi yang bising, seperti di tapak pembinaan, pertanian, kilang, perkhidmatan landskap, mencuci kereta, atau pekerjaan lain yang melibatkan peralatan atau mesin yang kuat? (Dengan pekerjaan bising, kami maksudkan bunyi yang begitu kuat sehingga anda perlu menjerit atau bercakap dengan suara kuat untuk didengari pada jarak yang dekat)	Adakah anda bekerja dalam keadaan terdedah kepada bunyi yang bising, seperti di tapak pembinaan, perladangan, kilang, perkhidmatan mesin rumput, mencuci kereta, atau pekerjaan lain yang melibatkan peralatan atau mesin yang kuat? (Dengan pekerjaan bising, kami maksudkan bunyi yang sangat kuat sehingga anda perlu menjerit atau bercakap dengan suara kuat untuk didengari pada jarak yang dekat)	pertanian	perladangan

Content Validation

For this study, the value of I-CVI for each questionnaire components are in the range of 0.87 to 1. Meanwhile, the S-CVI/Ave obtained is 0.98.

Table 4: I-CVI for 31 Items

Items	I-CVI
1a, 1b, 1c, 2a, 2b, 2c, 3a, 3b, 3c, 4a, 4b, 4c, 5a, 5b, 5c, 7a, 7b, 7c, 7d, 8a, 8b, 9a, 9b, 10a, 10b, 10c, 10d, 10e	1
6a, 6b, 6c	0.87

Table 5: Example of changes in semantics following content validators' feedback

Items	Changes	
	From	To
2a	Di luar waktu bekerja, berapa kerapkah anda memandu atau menggunakan jentera berat seperti traktor, trak, atau peralatan perladangan/penjagaan rumput seperti mesin pemotong rumput/pam daun?	Di luar waktu bekerja, berapa kerapkah anda memandu atau menggunakan jentera berat seperti traktor, trak, atau peralatan perladangan/penjagaan rumput seperti mesin pemotong rumput/peniup daun?

Face Validation

For this study, the value of I-FVI for each the questionnaire items are in the range of 0.9 to 1. Meanwhile, the S-FVI/Ave obtained is 0.97.

Table 6: I-FVI for 31 Items

Items	I-FVI
1b, 1c, 2a, 2b, 2c, 3a, 3b, 4a, 4b, 4c, 5a, 5b, 5c, 6a, 6c, 7d, 9a, 9b, 10a, 10b, 10c, 10d, 10e	1
1a, 3c, 6b, 7a, 7b, 7c, 8a, 8b	0.9

DISCUSSION

Adaptation Process

The adaptation process is a critical component of cross-cultural research, ensuring that a questionnaire or instrument remains relevant, understandable, and applicable to the target population (Beaton et al., 2000; Coskun Benlidayi & Gupta, 2024; Epstein et al., 2015). In this study, several changes were made to the original questionnaire to ensure that the items were culturally and contextually appropriate for the Malaysian population. The changes were primarily categorized as omission, rephrasing, and merging items, all of which were intended

to maintain the original intent of the questionnaire while making it more applicable to the local context.

In terms of omission, certain items were removed due to their lack of relevance or common usage in Malaysia. A notable example was the omission of the word "snowmobiles" in Item 4a, as this term is not commonly associated with Malaysian culture or geography. Other terms, such as motorcycles, jet skis, and speed boats, were retained as they are familiar and applicable to the Malaysian population. This adjustment ensured that the questionnaire did not include irrelevant items, which could lead to confusion or misunderstanding among respondents.

The rephrasing category involved altering the phrasing of questions to ensure they were both culturally appropriate and easier to understand. For instance, Items 10a and 11a, which initially distinguished between noisy paid jobs during and outside of summer, were rephrased and merged into a single item. This change was made because Malaysia does not have a seasonal summer break that is common in temperate climates, making the original distinction irrelevant. The revised item now asks about noisy paid jobs conducted in the past year, which includes work done during school holidays or free time, making the question more universally applicable across the year in the Malaysian context.

These changes were made with careful consideration of the cultural context and the goal of preserving the validity and reliability of the questionnaire. Ensuring that questions are both culturally relevant and linguistically clear is essential for maintaining the integrity of the data collected and ensuring that the responses are meaningful and comparable (Han et al., 2022; Olievera et al., 2022). By modifying the items without altering their original meaning, the adaptation process effectively enhanced the cultural sensitivity of the questionnaire while retaining its capacity to assess the intended constructs.

Translation Process

During the translation process, several modifications were made to ensure semantic, idiomatic, and cultural equivalence between the original and translated versions of the questionnaire. These changes were carefully documented in Table 3, following the expert reviews to address inconsistencies and improve clarity.

For instance, in Item 2a and 10a, the term "pertanian" (agriculture) was changed to "perladangan" (farming), as the latter more accurately represents the context of machinery use in the Malaysian cultural and occupational setting. While both terms are closely related, "perladangan" better aligns with local interpretations involving heavy-duty vehicles and land maintenance equipment. This change reflects the importance of selecting culturally appropriate terminology to enhance the relevance and comprehension of the questionnaire (McAllister et al., 2021; Sousa & Rojjanasrirat, 2011).

Similarly, for Item 3a, the phrase "dengan sistem pembesar suara" was restructured to "yang menggunakan sistem pembesar suara" to increase syntactic clarity and ensure the question flows naturally in Malay. Such grammatical realignments, although minor, are vital to preserve the intent of the original item while enhancing user comprehension (Abd Rahim et al., 2023; Gjersing et al., 2010).

Another significant adjustment was seen in Item 4a, where "motor bertayar empat" was replaced with the more precise "kenderaan empat roda" to prevent ambiguity. This change ensures a clearer understanding of the type of vehicle referenced. Thereby, reducing potential misinterpretation by respondents unfamiliar with technical phrasing.

Additionally, more familiar Malay terms like "fon kepala" and "fon telinga" replaced the English loanwords "headphone" and "earphone" (Items 8a and 9a). These adjustments helped preserve the original meaning while enhancing cultural relevance and respondent understanding.

Content Validation

Content validation was carried out by eight audiologists, with I-CVI values ranging from 0.87 to 1.00. Although three items related to firearm exposure (6a, 6b, and 6c) received I-CVI scores of 0.87, they remained above the acceptable threshold of 0.83 as recommended by Lynn (1986) and Yusoff (2019) for studies involving six or more experts. According to Zamanzadeh et al. (2015), items scoring between 0.70 and 0.79 require revision, while those below 0.70 should be eliminated. As all items scored above the threshold, none were removed from the questionnaire. Although no major revisions were required, minor semantic adjustments were made based on content validator feedback to enhance clarity and natural language flow. For example, the phrase "pam daun" was refined to "peniup daun" to better reflect standard Malay usage. The overall S-CVI/Ave score of 0.98 indicated excellent content validity, demonstrating that the items adequately represented the construct of noise exposure (Amy Sie-Yik et al., 2017; Polit & Beck, 2006).

Face Validation

For face validation, ten lay participants with diverse backgrounds assessed the clarity and comprehensibility of each item. The I-FVI scores ranged from 0.90 to 1.00, and the overall S-FVI/Ave was 0.97. Although a few items scored slightly lower, this still falls within acceptable levels as per Rahman et al. (2021). Most participants agreed that the questionnaire items were well-structured and easy to understand, indicating that the final version is suitable for self-administration.

Several limitations should be considered in interpreting the results of this study. First, the validation process focused mainly on content and face validity, without extending to other psychometric properties such as test-retest reliability, which would strengthen the robustness of the tool. Second, although expert and lay feedback were incorporated, the sample size for both groups was relatively small. A broader panel of experts and a larger, more diverse group of respondents would have provided richer insights and potentially different perspectives on item clarity and relevance. Lastly, the study was conducted in a controlled setting, which may not fully reflect how the questionnaire performs in real-world clinical or occupational screening environments.

CONCLUSION

The present study successfully translated, adapted, and validated the Malay version of the Noise Exposure Questionnaire (M-NEQ), demonstrating strong evidence of content and face validity. The translation process ensured both linguistic and cultural appropriateness, while expert and layperson evaluations confirmed the questionnaire's clarity and relevance. High I-CVI and S-CVI/Ave values support the relevance of each item, and similarly high I-FVI and S-FVI/Ave values indicate the tool is comprehensible to the intended population. These findings suggest that the M-NEQ is a valid instrument for assessing noise exposure in a Malaysian context. Further research is recommended to explore additional psychometric properties and assess its effectiveness in clinical and community settings.

ACKNOWLEDGEMENT

This study was conducted with the approval of the International Islamic University Malaysia (IIUM) Research Ethics Committee (IREC) and was not funded by any grant.

REFERENCES

Abd Rahim, M. H., Ibrahim, M. I., Ab Rahman, A., & Yaacob, N. M. (2023). Translation, Cross-Cultural Adaptation and Validation of Movement Behaviour Questionnaire into Malay Language (MBQ-M) for Measuring Movement Behaviors among Preschool Children in Kelantan, Malaysia. *Healthcare*, *11*(9), 1276. https://doi.org/10.3390/healthcare11091276

Amy Sie-Yik, L., Yusoff, M. S. B., Yeong-Yeh, L., Sy-Bing, C., Rashid, F., Wahid, N., Jin-Zhong, X., & Min-Tze, L. (2017). Development, Translation and Validation of Questionnaires for Diarrhoea and Respiratory-related Illnesses during Probiotic Administration in Children. *Education in Medicine Journal*, *9*(4), 19–30. https://doi.org/10.21315/eimj2017.9.4.3

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Coskun Benlidayi, I., & Gupta, L. (2024). Translation and Cross-Cultural Adaptation: A Critical Step in Multi-National Survey Studies. *Journal of Korean Medical Science*, *39*(49). https://doi.org/10.3346/jkms.2024.39.e336

Epstein, J., Santo, R. M., & Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of Clinical Epidemiology*, *68*(4), 435–441. https://doi.org/10.1016/j.jclinepi.2014.11.021

Gjersing, L., Caplehorn, J. R., & Clausen, T. (2010). Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Medical Research Methodology*, *10*(1), 13. https://doi.org/10.1186/1471-2288-10-13

Han, K., Wang, Q., Yang, L., Xu, S., Li, C., Lin, J., Wu, H., & Huang, Z. (2022). Development and initial validation of the Chinese Version of the Noise Exposure Questionnaire (C-NEQ). *BMC Public Health*, *22*(1), 1–7. https://doi.org/10.1186/S12889-022-12648-5/TABLES/5

Johnson, T. A., Cooper, S., Stamper, G. C., & Chertoff, M. (2017). Noise Exposure Questionnaire: A Tool for Quantifying Annual Noise Exposure. *Journal of the American Academy of Audiology*, *28*(01), 014–035. https://doi.org/10.3766/jaaa.15070

McAllister, A., Burström, B., & Corrigan, P. (2021). Cultural adaptation and validation of the Attribution Questionnaire for stigma towards disability pension applicants for use among psychiatrists and general

practitioners in Sweden. *BMC Psychology*, *9*(1), 27. https://doi.org/10.1186/s40359-021-00523-8

Neitzel, R. L., & Fligor, B. J. (2019). Risk of noise-induced hearing loss due to recreational sound: Review and recommendations. *The Journal of the Acoustical Society of America*, *146*(5), 3911–3921. https://doi.org/10.1121/1.5132287

Oliveira, L. C. de, Rocha, C. H., Matas, C. G., Paiva, K. M. de, Moreira, R. R., & Samelli, A. G. (2023). Translation and cross-cultural adaptation of the Noise Exposure Questionnaire (NEQ) to Brazilian Portuguese. *CoDAS*, *35*(3). https://doi.org/10.1590/2317-1782/20212022062en

Occupational Safety and Health (Noise Exposure) Regulations 2019.

Osmanoğlu, H., Dizdar, H. T., & Koçyiğit, A. A. (2024). The effects of music listening time with headphones on hearing thresholds among the young population. *The Egyptian Journal of Otolaryngology*, *40*(1), 13. https://doi.org/10.1186/s43163-024-00574-9

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Rahman, M. A., Yusoff, M. S. B., Roslan, N. S., Mohammad, J. A.-M., & Ahmad, A. (2021). Development and validation of the medical professionals resilience scale. *BMC Health Services Research*, *21*(1), 482. https://doi.org/10.1186/s12913-021-06542-w

Razali, A., Azmy, F. 'Izzaty, & Khairul Anwar, A. M. (2022). *Development and validation of Malaysian noise and chemical exposure questionnaire towards hearing among hospital workers - PubMed*. https://pubmed.ncbi.nlm.nih.gov/35902930/

Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, *17*(2), 268–274. https://doi.org/10.1111/j.1365-2753.2010.01434.x

Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). *Medknow*. https://doi.org/10.4103/sja.SJA_203_17

Wang, X., & Cheng, Z. (2020). Cross-Sectional Studies. *Chest*, *158*(1), S65–S71. https://doi.org/10.1016/j.chest.2020.03.012

World Health Organization (2024). Deafness and Hearing Loss. https://www.who.int/health-topics/hearing-loss#tab=tab_1

Yusoff, M. S. B. (2019). ABC of Content Validation and Content Validity Index Calculation. *Education in Medicine Journal*, *11*(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. *Journal of Caring Sciences*, *4*(2), 165–178. https://doi.org/10.15171/jcs.2015.017

ORIGINAL ARTICLE

Perceptions of Occupational Noise, Hearing Protection, and Risk Among Airport Workers: A Preliminary Study from Kuantan, Malaysia.

Noraidah Ismail, Corresponding Author[1] Ili Athirah Roseli[2], Nurul Syarida Mohd Sakeri[1]*

[1]Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2] Soundlife Healthcare (SA) Sdn Bhd., 39-1f, Jalan Setia Perdana BA U13/BA, Setia Alam, Seksyen U13, 40170 Shah Alam, Selangor, Malaysia

Keywords:
occupational noise, hearing protection, airport workers, noise-induced hearing loss, Malaysia, hearing conservation programs

ABSTRACT

Background: Occupational noise exposure remains a significant health concern in high-risk environments such as airports. High noise produced by an aircraft or its components, whether on the ground while parked, during take-off, or during landing, can cause problems for workers and nearby communities. Despite regulatory efforts, the use of hearing protection devices (HPDs) and awareness of noise-induced hearing loss (NIHL) vary across settings, particularly in regional airports.
Objective: This preliminary study aimed to explore airport workers' perceptions of occupational noise, identify perceived barriers to hearing protection, and examine readiness for noise-reduction behaviours among staff at Kuantan Airport, in Kuantan, Pahang, Malaysia.
Methods: A total of 29 Malaysia Airports Holdings Berhad (MAHB) employees participated, representing approximately 64% of the airport's workforce. Participants completed a self-administered questionnaire assessing attitudes toward noise, perceived susceptibility to NIHL, self-efficacy, perceived benefits of noise reduction, and perceived barriers to hearing protection. Hearing screening was also conducted using pure tone audiometry and tympanometry.
Results: Findings revealed high mean scores in the attitude, susceptibility, and barriers domains. Participants recognized the risks of occupational noise and expressed openness to protective behaviours. While discomfort and communication challenges were reported as barriers to HPD use, participants indicated that these challenges could be addressed through targeted interventions.
Conclusion: The results highlight a strong foundation for implementing effective hearing conservation programs in regional airports. Improving HPD design, delivering tailored education, and fostering workplace engagement can enhance protective behaviours and support long-term hearing health. Future studies with larger, more diverse samples are recommended to validate and extend these findings. These preliminary findings contribute to understanding NIHL risk perception and support the design of tailored conservation strategies in underserved aviation sectors.

INTRODUCTION

Occupational noise remains a major health and safety concern across various industries, with airport environments posing particularly high risks due to constant exposure to elevated noise levels from aircraft and ground operations (Bertsch & Sescu, 2022; Civil Aviation Authority, 2024; Nair & Kashyap, 2009; Jockel et al., 2019). Prolonged exposure to high noise has been associated with noise-induced hearing loss (NIHL), tinnitus, stress, sleep disturbances, and cardiovascular conditions (Abbasi et al., 2024; Chada et al., 2022; Hassan, 2024; Sivakumaran et al., 2022). According to Hailu et al. (2024), a 10 dB increase in aircraft noise exposure is linked to a 7% to 17% higher risk of cardiovascular disease.

Noise across various areas of an aircraft and airport can present significant hazards to workers' hearing, with exposure levels varying by role and location (Lindgren et al., 2009). Studies have recorded airport noise ranging from 70–91 dB(A), with peaks reaching approximately 119 dB(A) (Aziz et al., 2017; Civil Aviation Authority, 2024; Smedje et al., 2011). Cockpit noise typically ranges from 70–90 dB(A) (Zevitas et al., 2018), rising even higher during take-off and landing, potentially exceeding 140 dB(A) (Lee et al., 2022; Xu, 2025). Ground staff, maintenance crews, and baggage handlers are especially vulnerable, often exposed to levels exceeding occupational limits (Athirah & Shahida, 2019; Aziz et al., 2017; Hong & Kim, 2001; International Civil Aviation Organization [ICAO], 2013).

[1*] *Corresponding author.*
E-mail address: noraidah@iium.edu.my

Based on a study by Ozcan and Nemlioglu (2006), ground crews near taxiing aircraft were exposed to noise levels exceeding 140 dB(A), maintenance personnel near engines to 100–120 dB(A), and cabin staff to approximately 80–85 dB(A) during flights. The primary sources of these elevated levels are airframe and engine noise, supplemented by hydraulic, air conditioning, and communication systems (Xu, 2025). These findings highlight the urgent need for robust hearing conservation measures to protect the long-term occupational health of aviation personnel (Ivoševi ć et al., 2018; Orikpete et al., 2024).

While the physical hazards associated with occupational noise are well established, workers' perceptions of noise exposure remain a critical yet underexplored area. Despite the high risk of exposure, there is a notable lack of research examining airport workers' knowledge, perceptions, and awareness regarding occupational noise and its long-term effects. This knowledge is vital, as previous studies suggest that individuals who underestimate noise-related risks are less likely to use hearing protection devices (HPDs) or adhere to safety protocols. Conversely, heightened awareness and accurate risk perception are positively correlated with improved compliance (Guo et al., 2024; Ishak et al., 2023; Vosoughi et al., 2022). Understanding these perceptions is essential for developing targeted interventions and training programs that address the specific needs of this workforce, ultimately aiming to reduce NIHL and enhance the overall safety culture (Bandyopadhyay, 2024).

In Malaysia, where airports serve as key transport hubs, there is limited research on how airport workers perceive occupational noise risks. Although NIHL remains one of the most reported occupational illnesses (Aziz et al., 2017; Nasir & Rampal; 2012; Department of Statistics Malaysia [DOSM], 2023; Department of Safety and Health [DOSH], 2024), much of the existing literature focuses on large industrial sectors, overlooking smaller regional airports. Workers in these settings, such as Kuantan Airport in Pahang, may face unique challenges, including fewer resources for noise mitigation and limited enforcement of protective practices (Farah et al., 2017).

This preliminary study is part of a larger, ongoing project aimed at understanding and addressing occupational noise exposure among airport workers. Specifically, it explores airport staff's perceptions of occupational noise and document their self-reported hearing-related experiences and the use of hearing protection measures. By addressing this gap, the findings are expected to guide the development of targeted health and safety interventions, with a particular focus on regional airports where access to noise management resources is often limited.

METHODOLOGY

Study Design

This study employed a cross-sectional observational design to assess airport workers' perceptions, attitudes, and hearing health status related to occupational noise exposure. The design was chosen to capture a snapshot of workers' self-reported experiences and audiometric outcomes within their existing workplace environment.

Study Location

This investigation was carried out at Sultan Ahmad Shah Airport (Kuantan Airport) – located approximately 15 kilometres from the town centre of Kuantan, Pahang. As a dual-use facility operated by Malaysia Airports Holdings Berhad (MAHB), Kuantan Airport shares its runway with the Royal Malaysian Air Force (RMAF Kuantan), a configuration that contributes to consistently elevated occupational noise levels. The airport accommodates over one million passengers annually and maintains two daily commercial flights to and from Kuala Lumpur International Airport (PKNP, 2024).

Participants

Participants were recruited from the Malaysia Airports Holdings Berhad (MAHB) workforce at Kuantan Airport. A total of 29 employees volunteered to participate, representing approximately 64% of the airport's estimated 45-member MAHB staff. With institutional support from the MAHB Safety Office, participants were drawn from various units, including safety, technical and operations, engineering, sales, and cargo. All participants were Malaysian nationals aged between 33 and 58 years. Given the small and well-defined population, the high response rate supports the sample's representativeness for the scope of this study.

Questionnaire

To investigate workers' perceptions of occupational noise and self-reported hearing-related experiences, a two-phase data collection approach was adopted: (1) questionnaire administration and (2) audiometric screening. Prior to participation, informed consent was obtained from all respondents.

The questionnaire comprised two sections. The first section was on demographic and auditory health. This portion gathered socio demographic information (e.g., age, gender, education, job role) and included targeted

items about tinnitus, self-reported hearing difficulties, and perceived noise exposure. These questions allowed for capturing both environmental risk factors and participants' subjective auditory health status.

The second section deployed the Malay version of the Noise at Work (NAW) questionnaire originally developed by Purdy & Williams (2002) to evaluate attitudes and perceptions related to occupational noise exposure. This version, translated and psychometrically validated by Ismail (2013), demonstrated strong internal consistency in prior studies (Cronbach's α = .89–.93). In the current study, reliability analysis yielded a Cronbach's alpha of α = .87.

The NAW questionnaire comprised 20 items across five conceptual domains: (1) attitudes toward noise, (2) perceived barriers to hearing loss prevention, (3) perceived self-efficacy, (4) perceived benefits of noise reduction, and (5) perceived susceptibility to hearing loss. Items were rated on a 5-point Likert scale (1 = Strongly Agree to 5 = Strongly Disagree), with selected statements reverse scored to ensure interpretive consistency. Higher scores in the attitude and susceptibility domains were interpreted as indicating elevated concern about occupational noise and greater behavioural intent toward hearing conservation interventions.

Hearing Screening

The second phase of data collection involved pure tone audiometry (PTA) and tympanometry testing. All procedures were conducted in a sound-treated mobile unit stationed on the MAHB premises. Otoscopic examinations were performed using a Heine otoscope, while tympanometry was conducted with a Grason-Stadler GSI TympStar.

Hearing screening was performed using a calibrated portable Madsen audiometer, in accordance with ANSI S3.1-1999 standards. The minimum threshold was set at 20 dB HL for the 1, 2, and 4 kHz frequencies in each ear. Failure to respond at any of these frequencies constituted a refer/fail criterion. This threshold reflects established adult screening protocols and ensures consistency with standard practice, even in mobile settings. Ambient noise levels during testing were monitored using a sound level meter (SLM) and maintained within permissible limits for audiometric assessment in mobile environments.

In compliance with the Occupational Safety and Health (Noise Exposure) Regulations 2019, participants were required to abstain from significant noise exposure for at least 14 hours prior to testing. Individuals presenting with abnormal middle ear function—identified through Type B or Type C tympanograms—were referred to an audiology clinic for further diagnostic evaluation.

Procedures

Questionnaires were distributed and collected via an appointed officer at Kuantan Airport. Each participant received a comprehensive informed consent form attached to the questionnaire, which detailed the purpose and procedures of the study, and data confidentiality. Completed forms were returned in sealed envelopes to ensure confidentiality. Clear written instructions were provided to facilitate accurate and independent completion.

All data were systematically coded and analysed using SPSS software (Version 20).

RESULTS

Participant Characteristics

Of the 29 Malaysia Airports Holdings Berhad (MAHB) employees who participated, 95.7% were male. Ages ranged from 33 to 58 years (M = 44.6, SD = 7.62), and average work tenure was substantial (M = 22.4 years, SD = 8.89). Participants reported daily working hours of 8 to 12 hours (M = 9.13, SD = 1.25), with approximately two-thirds having attained secondary-level education. In terms of departmental affiliation, 30.4% were employed in operations, 26.1% in cargo, 13% in safety, and 4.3% in administration. Notably, 65.2% indicated prior exposure to informational material on NIHL.

These demographic findings suggest prolonged occupational exposure in a predominantly male workforce, which is consistent with trends in aviation ground operations and aligns with the study's focus on long-term auditory risk.

Self-Reported Hearing Difficulties and Noise Exposure

A significant proportion of participants (82.6%) self-reported hearing difficulties, and 69.6% indicated challenges hearing in noisy environments. Familial concern regarding hearing health was present for 56.5% of respondents, while 87% reported having undergone previous hearing assessments. Tinnitus emerged as a prevalent symptom, reported occasionally by 78.3%, frequently by 17.4%, and constantly by 4.3% of participants.

Regarding workplace noise exposure, 61% identified high levels of acoustic stress, and 59.1% indicated use of

hearing protection devices. Additionally, 60.9% reported never having smoked—a relevant factor when considering hearing health and susceptibility. These findings reflect high self-awareness of hearing impairment and notable prevalence of auditory symptoms, consistent with chronic exposure patterns identified in airport environments.

Objective Hearing Screening Outcomes

Of the 29 participants, nine underwent hearing screening through audiometric testing. Limited participation was attributed to scheduling challenges and operational demands. Among these, eight (88.9%) failed to meet minimum hearing threshold criteria and were referred for further diagnostic evaluation.

The comparison between self-reported hearing difficulties, family concern, and audiometric outcomes is summarized in **Table 1**. As shown, six of the seven individuals who self-reported hearing difficulties also failed the screening. Notably, two participants who denied having hearing problems likewise failed, suggesting a potential gap between perceived and actual hearing status. Similarly, three of the four participants whose family members had expressed concern about their hearing failed screening, while five without such concern also did not pass. These discrepancies highlight the limited accuracy of self-assessment and underscore the importance of routine clinical evaluation to complement personal awareness

Table 1: Hearing Screening and Self-Reported Hearing (N=9)

Subject	Self-Reported Hearing Problem	Family Concern	Hearing Screening Result
1	Yes	Yes	Fail
2	Yes	Yes	Fail
3	Yes	No	Fail
4	Yes	No	Fail
5	Yes	No	Fail
6	Yes	Yes	Fail
7	No	No	Fail
8	No	No	Fail
9	Yes	No	Fail

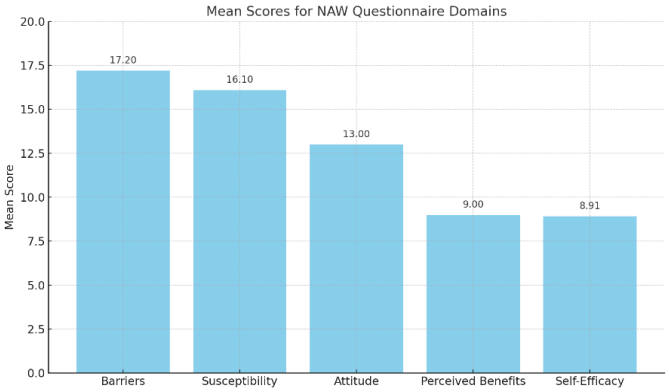
Note. All nine participants underwent audiometric screening. "Fail" indicates hearing thresholds above 20 dB HL at one or more test frequencies (1, 2, or 4 kHz) in either ear.

Noise at Work (NAW) Questionnaire Analysis

Mean scores across the five domains of the NAW questionnaire are presented in Figure 1. The highest values were observed in the Perceived Barriers (M = 17.2) and Perceived Susceptibility (M = 16.1) domains, followed by Attitude Toward Noise (M = 13.0), Perceived Benefits of Prevention (M = 9.0), and Self-Efficacy (M = 8.91). Higher scores reflect stronger perceptions of risk and the importance of noise reduction.

These findings suggest that while participants recognized the risks associated with occupational noise and expressed annoyance toward it, internalized confidence in managing or preventing hearing loss (self-efficacy) was comparatively low.

Figure 1. Mean Scores Across NAW Questionnaire Domains



DISCUSSION

This study examined perceptions of occupational noise and hearing health among personnel at Kuantan Airport—a regional facility noted for persistently high ambient noise levels associated with aviation operations. Analysis of self-reported data and audiometric screening revealed a high prevalence of hearing difficulties and tinnitus symptoms, alongside an elevated rate of screening failures. These outcomes reflect the cumulative impact of chronic noise exposure and highlight the need for comprehensive hearing conservation programs in comparable aviation settings.

Participants' responses to the Noise at Work (NAW) questionnaire revealed high perceived susceptibility and negative attitudes toward workplace noise, indicating general awareness of associated health risks. Notably, the "Perceived Barriers" domain received the highest scores, pointing to discomfort, communication challenges, and inconsistent access to hearing protection as key impediments to preventative action. These findings align

with prior research suggesting that such barriers may undermine workers' confidence in managing noise risks and hinder compliance with safety measures (Doutres et al., 2022; Ishak et al., 2023; Sameli et al., 2017).

Despite widespread awareness, low scores in self-efficacy and perceived benefits suggest limited confidence in adopting protective behaviours. Similar trends have been documented in other occupational contexts, emphasizing the importance of perceived control and risk perception in shaping hearing conservation outcomes (Fauzan et al., 2023; Orikpete et al., 2024; Vosoughi et al., 2022). These findings indicate that awareness alone may be insufficient to drive sustained behavioural change without targeted, supportive interventions. Practical strategies—such as improving the comfort and usability of hearing protection devices (HPDs), increasing access to routine audiometric evaluations, and enhancing communication during training—are essential to strengthen self-efficacy and facilitate consistent protective behaviour (Sedghi Noushabadi, 2024).

The discrepancy between self-assessed hearing and audiometric results is particularly concerning. Several participants who did not report hearing problems nevertheless failed screening, revealing gaps in self-awareness. This underscores the need to pair educational initiatives with clinical surveillance to promote early intervention. Periodic hearing assessments, combined with personalized risk communication, can bridge this perception gap and foster a proactive approach to occupational hearing health.

However, only a subset of participants (9 out of 29) completed the audiometric screening due to scheduling constraints and operational duties, which limited the strength of the objective findings. Future research could improve participation in hearing assessments by offering more flexible screening schedules, integrating testing into regular occupational health routines, or providing small incentives or honoraria to encourage employee engagement. These adjustments would enhance data completeness and strengthen the validity of future analyses.

Additionally, the study highlights the importance of context-specific solutions for regional airports such as Kuantan, where resource limitations and infrastructure constraints challenge the implementation of standard conservation measures. These findings support broader efforts to develop scalable, evidence-based interventions tailored to aviation environments and other high-noise industries. Sustained initiatives—both technical and behavioural—are critical to improving long-term occupational safety (Doutres et al., 2022; Fauzan et al., 2023; Ismail, 2013).

While the single-site design and modest sample size limit the generalizability of the findings, the high participation rate strengthens their internal validity. Future research should expand to include a larger sample from multiple airports and diverse occupational roles to enable comparative analysis and enhance external validity. Incorporating advanced psychometric tools may also help capture behavioural determinants more comprehensively. Such efforts will be instrumental in shaping national policy and targeted health programs for Malaysia's evolving aviation workforce.

CONCLUSION

This preliminary study offers valuable insights into airport workers' perceptions of occupational noise and their attitudes toward hearing conservation practices within a regional aviation context. Findings from Kuantan Airport revealed widespread self-reported hearing difficulties, frequent tinnitus symptoms, and a high failure rate in audiometric screening—underscoring the long-term health risks associated with chronic noise exposure. Participants demonstrated strong awareness of NIHL and expressed openness to protective behaviour, yet low self-efficacy and perceived benefits signal a disconnect between risk recognition and sustained action.

Importantly, barriers such as discomfort with HPDs, communication interference, and limited access to protective resources were prominent but not perceived as insurmountable. Addressing these challenges through regular hearing evaluations, improved HPD design, and tailored training programs could significantly empower workers and reinforce protective behaviours. Priority should be given to implementing targeted noise awareness campaigns and toolbox talks for airside ground crew and maintenance staff, integrating hearing protection modules into regular aviation safety briefings, and mandating annual audiometric screening within the airport's occupational health surveillance. These findings highlight the importance of combining education with workplace-specific interventions to promote meaningful and sustained behavioural change.

Although this study was conducted at a single regional airport with a modest sample size, the high participation rate enhances the internal validity of its findings. The results provide an important foundation for developing scalable, context-specific hearing conservation strategies applicable to Malaysia's regional airports and other high-noise occupational settings. Future research should

incorporate larger and more diverse samples while exploring behavioural and organisational enablers of long-term compliance to support national policy development and strengthen occupational health frameworks.

From a practical standpoint, policymakers and airport management should prioritize the institutionalization of hearing conservation within workplace safety systems. This includes enforcing existing noise exposure regulations, implementing routine hearing surveillance, improving access to comfortable and effective hearing protection devices, and delivering regular awareness and training programs tailored to operational realities. Continuous evaluation of these initiatives will be essential to ensure sustained effectiveness and foster a culture of proactive hearing health among aviation personnel.

ACKNOWLEDGMENTS

This study was conducted with the approval of the International Islamic University Malaysia (IIUM) Research Ethics Committee (IREC) and was not funded by any grant.

REFERENCES

Abbasi, M., Derakhshan, J., & Darabi, F. (2024). The impact of noise-induced hearing loss on individual job performance: Exploring the role of aggression and work-related quality of life. *BMC Psychology, 12*(1), Article 113. https://doi.org/10.1186/s40359-024-02113-w

Athirah, B., & Nurul Shahida, M. (2019). Occupational noise exposure among airport workers in Malaysia: An ergonomic investigation. *Journal of Physics: Conference Series, 1262*, 012023. https://doi.org/10.1088/1742-6596/1262/1/012023

Aziz, A. R. A., Abd Aziz, S., & Hassan, M. Z. (2017). Noise exposure among maintenance crews of C130H aircraft lead to the hearing impairment. *Malaysian Journal of Public Health Medicine, 1*, 89–94.

Bertsch, L., & Sescu, A. (2022). *Aircraft noise*. MDPI. https://elib.dlr.de/187954/1/ELIB-Eintrag-2022-BertschL-187954-Sammelband.pdf

Bandyopadhyay, A., Mukherjee, A., Dhar, G., & Rout, A. J. (2024). Burden and risk factors of occupational noise-induced hearing loss among employees working at an international airport in Eastern India. *Indian Journal of Occupational and Environmental Medicine, 28*(3), 223–227. https://doi.org/10.4103/ijoem.ijoem_163_23

Chadha, S., Kamenov, K., & Cieza, A. (2021). The world report on hearing, 2021. *Bulletin of the World Health Organization, 99*(4), 242–242A. https://doi.org/10.2471/BLT.21.285643

Civil Aviation Authority. (2024). *Aviation noise and health*. https://www.caa.co.uk/passengers-and-public/environment/noise/aviation-noise-and-health/

Doutres, O., Terroir, J., Jolly, C., Gauvin, C., Martin, L., & Negrini, A. (2022). Towards a holistic model explaining hearing protection device use among workers. *International Journal of Environmental Research and Public Health, 19*(9), 5578. https://doi.org/10.3390/ijerph19095578

Department of Statistics Malaysia (DOSM). (2023). *National occupational injury 2023*. Institute of Labour Market Information and Analysis (ILMIA). https://www.ilmia.gov.my/index.php/en/bda-noa

Department of Safety and Health (DOSH). (2024). *Occupational diseases and poisoning statistics*. https://www.dosh.gov.my/index.php/statistic-v/occupational-diseases-and-poisoning-statistics

Farah, A., Ahmad, R., & Ramli, R. (2017). Occupational noise exposure among airport ground staff in Malaysia. *Malaysian Journal of Medicine and Health Sciences, 13*(2), 36–42.

Fauzan, N. S., Sukadarin, E. H., Widia, M., Irianto, I., & Ghazali, I. (2023). A systematic literature review of the factors influencing hearing protection device usage among industrial workers. *International Journal of Environmental Research and Public Health, 20*(4), 2934. https://doi.org/10.3390/ijerph20042934

Feder, K., Michaud, D., McNamee, J., Fitzpatrick, E., Davies, H., & Leroux, T. (2017). Prevalence of hazardous occupational noise exposure, hearing loss, and hearing protection usage among a representative sample of working Canadians. *Journal of Occupational and Environmental Medicine, 59*(1), 92–113.

Guo, J., Shu, L., Wen, W., Xu, G., Zhan, L., Yan, M., Long, T., Fan, Z., Wu, J., & Xiao, B. (2024). The influencing factors of hearing protection device usage among noise-exposed workers in Guangdong Province: A structural equation modeling-based survey. *BMC Public Health, 24*, 1044. https://doi.org/10.1186/s12889-024-18428-7

Hailu, A., Zeleke, B., Ermias, Z., et al. (2024). Prevalence

and associated factors of noise-induced hearing loss among workers in Bishoftu Central Air Base of Ethiopia. *Scientific Reports, 14*, 10762. https://doi.org/10.1038/s41598-024-56977-4

Hassan, N. (2024, November 8). Noise pollution and its effects on human health: A review. *EPRA International Journal of Multidisciplinary Research (IJMR), 10*, 24–37. https://doi.org/10.36713/epra18872

Hong, O. S., & Kim, M. J. (2001). Factors associated with hearing loss among workers of the airline industry in Korea. *ORL–Head and Neck Nursing, 19*, 7–13.

International Civil Aviation Organization. (2013). *Airport services manual Part 2: Pavement design and evaluation (Doc 9157)*. ICAO.

Ishak, N. A., Fauzan, N. S., & Widia, M. (2023). Correlation between noise-related risk perception, knowledge, and the use of hearing protection device: A case study among landfill workers. *AIP Conference Proceedings, 2682*(1), 060010. https://doi.org/10.1063/5.0114277

Ismail, N. (2013). *The effectiveness of occupational noise management in Malaysia* [Doctoral dissertation, Macquarie University]. Figshare. https://doi.org/10.25949/19433759.v1

Ivošević, J., Bucak, T., & Andraši, P. (2018). Effects of interior aircraft noise on pilot performance. *Applied Acoustics, 139*, 8–13. https://doi.org/10.1016/j.apacoust.2018.04.006

Jokel, C., Yankaskas, K., & Robinette, M. B. (2019). Noise of military weapons, ground vehicles, planes and ships. *The Journal of the Acoustical Society of America, 146*(5), 3832–3838. https://doi.org/10.1121/1.5134069

Lee, H. P., Kumar, S., Garg, S., & Lim, K. M. (2022). Assessment of in-cabin noise of wide-body aircrafts. *Applied Acoustics, 194*, 108809. https://doi.org/10.1016/j.apacoust.2022.108809

Lindgren, T., Wieslander, G., Nordquist, T., Dammström, B. G., & Norbäck, D. (2009). Hearing status among cabin crew in a Swedish commercial airline company. *International Archives of Occupational and Environmental Health, 82*(7), 887–892. https://doi.org/10.1007/s00420-008-0372-7

Nair, S., & Kashyap, R. (2009). Prevalence of noise induced hearing loss in Indian air force personnel. *Medical Journal of the Armed Forces India, 65*(3), 247–251.

Nasir, H., & Rampal, K. (2012). Hearing loss and contributing factors among airport workers in Malaysia. *The Medical Journal of Malaysia, 67*(1), 81–86.

Occupational Safety and Health in Malaysia Annual Report.

Occupational Safety and Health Administration. (2020). *Protecting workers from noise hazards in the workplace*. U.S. Department of Labor.

Orikpete, O. F., Dennis, N. M., Kikanme, K. N., & Ewim, D. R. E. (2024). *Advancing noise management in aviation: Strategic approaches for preventing noise-induced hearing loss*. Journal of Environmental Management, *363*, Article 121413. https://doi.org/10.1016/j.jenvman.2024.121413

Ozcan, H. K., & Nemlioglu, S. (2006). In-cabin noise levels during commercial aircraft flights. *Canadian Acoustics, 34*(4), 31–35. https://jcaa.caa-aca.ca/index.php/jcaa/article/view/1854

PKNP. (2024). *The official portal: Invest in Pahang*. https://investinpahang.gov.my/invest-in-pahang/infrastructure/the-sultan-ahmad-shah-airport/

Purdy, S. C., & Williams, W. (2002). Development of the Noise at Work Questionnaire to assess perceptions of noise in the workplace. *Journal of Occupational Health Safety–Aust NZ, 18*(1), 77-83.

Samelli, A. G., Gomes, R. F., Chammas, T. V., Silva, B. G., & Moreira, R. R. (2018). The study of attenuation levels and the comfort of earplugs. *Noise & Health, 20*(1), 112–118.

Sivakumaran, K., Ritonja, J. A., Waseem, H., AlShenaibar, L., Morgan, E., Ahmadi, S. A., Denning, A., Michaud, D. S., & Morgan, R. L. (2022). Impact of noise exposure on risk of developing stress-related health effects related to the cardiovascular system: A systematic review and meta-analysis. *Noise & Health, 24*(114), 107–129. https://doi.org/10.4103/nah.nah_83_21

Sedghi Noushabadi, Z. (2024). The influence of risk perception on hearing conservation and the use of hearing protection devices. IntechOpen. https://doi.org/10.5772/intechopen.1007120

Smedje, G., Lunden, M., Gartner, L., Lundgren, H., & Lindgren, T. (2011). Hearing status among aircraft maintenance personnel in a commercial airline

company. *Noise & Health, 13*(54), 364–370.

Xu, Y. (2025, January 13). Reviewing of research on aircraft noise reduction technology. *Applied and Computational Engineering, 130*, 152–158. https://doi.org/10.54254/2755-2721/2025.20307

Vosoughi, S., Alimohammadi, I., Hosseini, A. F., & Sedghy, Z. (2022). Knowledge, attitude, and practice (KAP) regarding hearing protection: A case study among offset printing industry workers in Tehran province. *Noise Control Engineering Journal, 70*(3), 270–278. https://doi.org/10.3397/1/377021

Zakia, M. A., Jani, J., & Azman, S. (2024). The impact of aircraft noise on residents residing in the vicinity of Subang Airport. *Jurnal Kejuruteraan, 36*(5), 2155–2164. https://doi.org/10.17576/jkukm-2024-36(5)-33

Zevitas, C. D., Spengler, J. D., Jones, B., McNeely, E., Coull, B., Cao, X., et al. (2018). Assessment of noise in the airplane cabin environment. *Journal of Exposure Science & Environmental Epidemiology, 28*(6), 568–578. https://doi.org/10.1038/s41370-018-0027-z

ORIGINAL ARTICLE

Validation of Malay Version of the Parents' Feelings and Perceptions Towards Their Child's Speech or Language Disorder (M-PFPSLD) Questionnaire

Amni Syamimi Mohamad Khairi[1], Masnira Jusoh[2,*]

[1]Footprints Occupational Therapy Center, Lot A77, Jalan 2 Cinta Sayang Resort Home, 08000 Sungai Petani, Kedah

[2]Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota 25200 Kuantan, Pahang, Malaysia

Keywords:
parents' feelings; parents' perception; speech or language disorder; validation; psychometric properties

ABSTRACT

Background: There is a limited number of Malay versions of questionnaires that assess the parents' feelings and perceptions towards their child diagnosed with a speech or language disorder. Hence, research on the validation of such a survey is necessary to expand the range of assessment and screening tools available in Malay for use within the Malay-speaking population. Objective: The aims of the current study are: to examine the content validity of the M-PFPSLD questionnaire; to investigate the face validity of the M-PFPSLD questionnaire; and to determine the appropriateness of the grammatical and sentence structures of the M-PFPSLD questionnaire. **Methods:** The M-PFPSLD questionnaire was validated using a cross-cultural adaptation and validation process that includes the process of 1) expert committee review; 2) face validation and pre-testing among parents with typically developing children; and 3) proofreading. **Results:** The item-level CVI (I-CVI) of 27 out of 28 items were 0.80 and above, which indicated the appropriateness of the items. For scale-level CVI (S-CVI), the average S-CVI (S-CVI/Ave) of the second version of the harmonised M-PFPSLD questionnaire (BM-H2) was 0.93, while the S-CVI universal agreement (S-CVI/UA) was 0.71. All 28 items in the BM-H2 version were acceptable with a 100% score of face validity index (FVI) based on its criteria of readability, feasibility, clarity of words, and layout and style. No changes were made after the proofreading process, thus indicating that the BM-H2 version was well-written, grammatical-error-free, and accurately expressed the intended meaning of the original context. **Conclusion:** This study concluded that the M-PFPSLD questionnaire has a good content validity index (CVI) and face validity index (FVI). Therefore, it shows that the findings may act as preliminary psychometric properties of the M-PFPSLD questionnaire, and future studies can further research the psychometric properties of the M-PFPSLD questionnaire, such as reliability, construct validity, convergent and divergent validity.

INTRODUCTION

American Speech-Language-Hearing Association (ASHA) (2018) defined speech or language disorder as the difficulty in articulating speech sounds, comprehending language, or producing spoken or written language. Communication is required for daily functioning, such as relaying information, requesting needs, expressing emotions and feelings, and establishing relationships (Kiogora, 2021). Children presented with speech or language disorders may be impacted significantly which reduces their quality of life (QOL) as these children are at higher risk for socioemotional difficulties such as socialising and internalising compared to the typically developing children (Foster, Choo, & Smith, 2023). Children's communication problems may affect not only their educational, social skills, and emotional development, but also will be major challenges to their parents as communication is strongly associated with improving the QOL (Lopez et al., 2021). Parental experiences may be overlooked when they spend most of their time managing children with speech and language disorders. Parents of children with speech or language disorders may face many challenges in the aspects of their child's social skills, communication skills, academic performance, and emotional development (Lopez et al. 2021). Emotional factors are crucial as parents get involved directly with their child's speech and language development (Aliza & Umithayyibah, 2021). Parents may have concerns regarding their child's condition and how it will affect the future of their children. Parents may face more challenges if their children have other medical conditions other than speech and language disorders such as autism spectrum disorder (ASD) as there is evidence

* *Corresponding author.*
E-mail address: masnirajusoh@iium.edu.my

that speech or language disorders are usually linked with other health problems or medical conditions such as ASD, attention deficit hyperactivity disorder (ADHD), down syndrome, cerebral palsy, and low birth weight (Norman, 2021).

It is important for the parents of children with speech or language disorders to be fully aware and understand the children's impairments and the effects on the children's future if lack of service provision and late intervention. McGregor (2020) also described that parents of children with speech or language disorders may find the terminology used in therapy sessions such as developmental language disorder (DLD) or expressive language delay, unfamiliar, and that language impairments can be less visible without other medical diagnosis, such as ASD or Down Syndrome (cited in Lopez et al., 2021). A lack of awareness and visibility of speech or language impairments compared to when the impairments are linked with other medical conditions, such as Down Syndrome, may lead to a lack of service provision and knowledge of the children's speech or language disorders, which should be helpful for the parents to understand their child's disorder better (Lopez et al., 2021). Parents spend most of their time with their children at home compared to professionals. Thus, parents should be actively involved in speech-language activities at home, as the time allocated to doing speech-language activities with children will help in the child's progress (Aliza & Umithayyibah, 2021).

Studying the parents' perspectives on their child's disorder will not only provide insights to the professionals but also increase the awareness and knowledge of the child's disorders among parents, which will lead to the satisfaction of parents, children, and professionals when there is an improvement shown in the children's speech and language abilities. The Parents' Perceptions and Feelings towards their Child's Speech or Language Disorders (PFPSLD) questionnaire was selected because this questionnaire specifically explores parents' emotional and perceptual responses towards their child's speech, language, or communication difficulties, as well as coping strategies related to their child's speech or language disorder. This questionnaire was developed to capture the unique parental experiences associated with communication difficulties rather than general developmental conditions (Holley, 2018). In comparison, several other questionnaires also measure related but in broader constructs. For instance, Parenting Stress Index (PSI-4; Abidin, 2012) focuses mainly on general parenting stress and parent-child interaction, rather than perceptions about communication difficulties, Family Impact Questionnaire (FIQ; Donenberg & Baker, 1993) examines the impact of a child's disability on family functioning, but does not address specific feelings or perceptions towards speech or language difficulties, and PedsQL Family Impact Module (Varni et al., 2004) evaluates parental QOL and family functioning. Still, it is not tailored for parents of children with communication difficulties. The PFPSLD questionnaire covers a range of topics associated with a child's communication problem including questions that reflect the overall quality of life (QOL) and well-being of the child, aligning with the holistic framework of family-centered speech-language intervention (Holley, 2018). This questionnaire is also relevant for clinical practice as it was originally developed for use by speech-language therapists working closely with families of children with communication disorders. Therefore, the PFPSLD questionnaire is more suitable for this study as it directly explores parents' feelings and perceptions related to their child's speech or language difficulties.

However, the original version of the PFPSLD questionnaire was developed in English and is only suitable for the English-speaking population (Holley, 2018). According to Holley (2018), no published studies have reported translation or validation of PFPSLD questionnaire into other languages which supports the importance of validating a Malay version of PFPSLD to ensure cultural and linguistic relevance for the Malay-speaking population. Thus, the information regarding the feelings and perceptions of parents of children with speech or language disorders in various nations was limited, especially in non-English speaking countries where language screening and language disorder services are provided (Lopez et al., 2021). The Malay version of the parents' perceptions and feelings towards their child's speech or language disorders (M-PFPSLD) questionnaire aims to examine how parents in Malaysia construed and described the speech and language disorders of their child. If the original English version of PFPSLD is distributed to the Malay-speaking population, this will influence the way services will be provided, knowing that there will be some cultural similarities and differences in the population (Squires et al., 2020, as cited in Lopez et al., 2021).

Barratt, Shangase, and Msimang (2012) highlighted the necessity to be cautious in using translated materials that are not appropriately and systematically adapted for local usage to meet the patient's needs and achieve the patient's satisfaction. Silva, Lamonica, and Hage (2021) also mentioned in their study the need to consider the cultural factors and the language spoken in a social setting. It is also important for the translated word or sentence to accurately express the intended meaning in the target population's language and culture (Silva, Lamonica, &

Hage, 2021). Therefore, validating the translated and adapted version of the M-PFPSLD questionnaire is essential to match the original purpose of the questionnaire in the Malay-speaking population. The aims of the current study are to examine the content validity of the M-PFPSLD questionnaire, to investigate the face validity of the M-PFPSLD questionnaire, and to determine the appropriateness of the grammatical and sentence structures of the M-PFPSLD questionnaire.

MATERIALS AND METHODS

Materials

This study was conducted using the Parents' Feelings and Perceptions Towards Their Child's Speech or Language Disorder (PFPSLD) questionnaire (Holley, 2018) (please refer to Appendix A) and the translated and adapted Malay version of the PFPSLD (M-PFPSLD) questionnaire (Jamaludin & Jusoh, 2023) (please refer to Appendix B). The M-PFPSLD questionnaire includes questions that reflect the child's general well-being and overall quality of life. There is a total of 28 questions that cover a range of topics relating to a child's communication impairment. To address the following two primary questions, which are the child's general well-being and overall quality of life, two main domains were divided: (1) the parents' feelings towards their child's speech or language disorder; and (2) the parents' perceptions regarding the impact of their child's disorder on their own lives (Holley, 2018). Responses are recorded using a Likert scale, which ranges from "Strongly Agree" (1 point) to "Strongly Disagree" (5 point). Most items are negatively worded; therefore, higher scores reflect more positive parental perceptions and lower scores indicate greater perceived challenges among parents of children with speech or language disorders. Through this questionnaire, parents can reflect on their feelings and opinions about speech or language disorders.

Subjects

The subjects were determined based on the stages of the study: i) content validation, ii) face validity and pre-testing, and iii) proofreading. In total, six expert committees with four different academic backgrounds, specifically, two speech-language therapists, two audiologists, a psychiatrist, and a clinical psychologist, were involved in the content validation process. Eight parents who have a child below 18 years old were involved in the face validity and pre-testing process of the initial version of the M-PFPSLD questionnaire. According to Yusoff (2019) and Burn (1995), the face validation and pre-testing process should involve approximately eight to eleven non-expert members to ensure sufficient feedback and variability in responses for evaluating content clarity. Therefore, eight to eleven parents were included in this process to support the validity of the assessment tool. As for the proofreading process, two Malay language experts, specifically, two Bahasa Melayu (BM) teachers were involved.

Cross-Cultural Adaptation and Validation Procedures

This study was approved by the International Islamic University Malaysia (IIUM) Research and Ethics Committee (IREC) before the study was conducted (IREC. No: KAHS 108/23).

The aim of this M-PFPSLD adaptation and validation was to produce a Malay version of the PFPSLD (M-PFPSLD) questionnaire that was conceptually and contextually equivalent to the original version and was suitable to be used by the Malay-speaking parents, teachers, and related professionals. Therefore, the M-PFPSLD questionnaire was validated using a cross-cultural adaptation and validation process in accordance with the guidelines by Beaton et al. (2000) and the World Health Organization [WHO] (2017).

The validation process that was conducted consisted of three stages and was summarised in Figure 1. The details of the validation procedures were further discussed in the following sections.

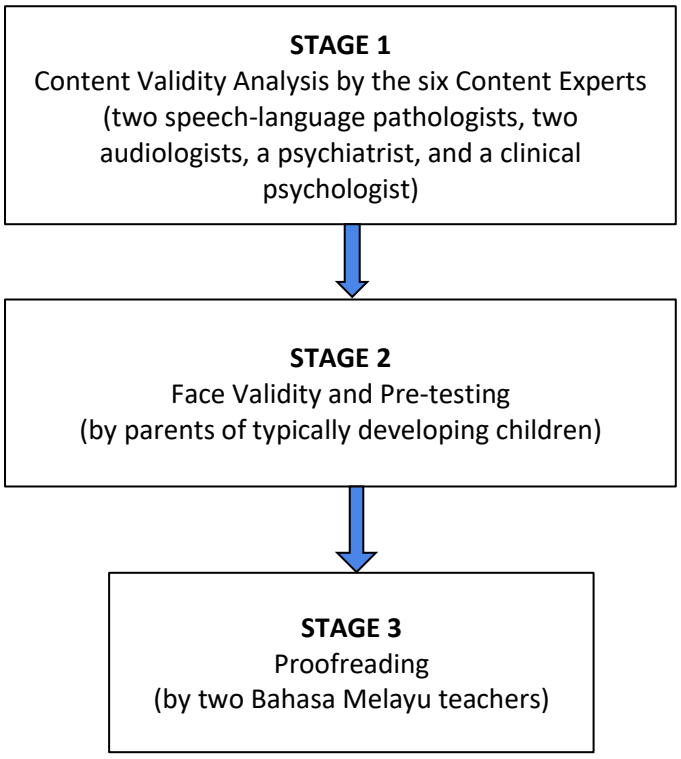


Figure 1: Flowchart of the Validation Procedures for the M-PFPSLD Questionnaire

Stage 1: Content Validation Procedures

The harmonised Malay PFPSLD (BM-H) and harmonised English PFPSLD (E-H) together with the original version of the questionnaire were reviewed by an expert committee consisting of six members (n=6) from four different academic backgrounds, specifically, two backgrounds in speech-language pathology, two in audiology, one in psychiatry, and one in psychology. The committee members consisted of multidisciplinary backgrounds, which was in line with the recommendation by Guillemin et al. (1993), which stated that an expert committee should consist of experts in the disease explored and the concept to be explored. In this study, the six expert committee members were recruited based on their professional expertise and experience in communication and language disorders. Invitations were sent via email to professionals working in universities and clinical settings who met the inclusion criteria of having at least five years of clinical or research experience related to speech and language disorders. Those who agreed to participate were provided with appointment letters as a content expert before the review process. There were two main aspects discussed during this process, which were the overall review by the expert committee and the content validity analysis of the questionnaire (Yusoff, 2019). For the overall review, each content expert reviewed the translated versions of the questionnaire to identify any discrepancies and inconsistencies in the translated versions compared to the original version. The content validity analysis was measured qualitatively and quantitatively.

A qualitative content validity analysis was carried out through the expert committees to review the consistency and differences between the original and the translated versions of the questionnaires in terms of wording, grammar, sentence structure, and scoring (Safikhani et al., 2013). Comments and suggestions from the expert committees were discussed, and amendments were made after the review to produce a finalised second version of the Harmonised BM-PFPSLD (BM-H2).

The BM-H2 version was measured quantitatively through quantitative content validity analysis. Two types of Content Validity Index (CVI), which are the Item-level CVI (I-CVI) and the overall scale-level CVI (S-CVI), were used in the study. To determine I-CVI, the expert committees were requested to rate the items in terms of intelligibility and relevancy to the underlying concept, using a 4-point Likert scale. The range of the Likert scale was from not relevant (1 point) to highly relevant (4 points) (Davis, 1992). Then, the I-CVI was computed based on the experts' agreement (with a score of 3 or 4 points given by the expert committees for each item) divided by the total number of content experts. The I-CVI must be equal to or higher than 79% to be considered appropriate (Abdullah et al., 2010). Apart from I-CVI, S-CVI was also computed. Pilot et al. (2007) mentioned that a minimum score of 0.80 for both S-CVI/Ave and S-CVI/UA was considered acceptable, while a score of 0.90 and above would be considered excellent (cited in Haidari and Uzun, 2019).

Stage 2: Face Validity and Pre-Testing Procedures

Face validity and pretesting was conducted by requesting the parents with typically developing children to read the questionnaire and provide feedback by stating 'Yes' or 'No' in terms of the readability, feasibility, clarity of words, and layout and style. They were approached personally or via WhatsApp using a viral message which included a brief explanation about the study. At this stage, it was not necessary to include parents of children with speech or language disorders, as the aim was to ensure that all items in BM-H2 version were easily understood in terms of the clarity and comprehension by the general parent population. This approach is consistent with the recommended practices in questionnaire development and adaptation (Boateng et al., 2018). The percentage of the face validity index (FVI) was calculated by dividing the raters in agreement and the total number of raters (Yusoff, 2019). For example, the FVI of question 1 is 9 divided by 10 raters which is equal to 0.9. According to Yusoff (2019), a percentage of 90% and above is considered acceptable. The participants were requested to answer the questionnaires, state the time taken to do so, and evaluate the items.

Stage 3: Proofreading Procedures

The BM-H2 version of the questionnaire was proofread by two Malay language experts. Both experts had been working as Bahasa Melayu teachers and were recruited based on their teaching experience in Bahasa Melayu for at least five years. Both Bahasa Melayu teachers were contacted through professional email invitations and voluntarily agreed to review the questionnaire. Both experts reviewed the wording, grammar, and sentence structure in the BM-H2 version.

Data Analysis

The data analysis for content validation, face validity and pre-testing, and proofreading process were summarised based on the specific objectives as shown in Table 1.

Table 1: Summary of Outcome and Data Analysis

Objectives	Outcomes	Data Analysis
To examine the content validity of the M-PFPSLD questionnaire.	Comments and suggestions from the expert committees	Qualitative content validity analysis
	Scores of each item	Quantitative content validity analysis - CVI (I-CVI and S-CVI) I-CVI - equal or higher than 79% was considered as appropriate (Abdullah et al., 2010) S-CVI - A minimum score of 0.80 was considered acceptable, while a score of 0.90 and above would be considered excellent (Pilot et al., 2007 cited in Haidari and Uzun, 2019)
To investigate the face validity of the M-PFPSLD questionnaire.	Scores of feasibility, comprehensibility, clarity, and consistency of each item	Face validity index (FVI) – percentage (90% and above should be considered acceptable) (Yusoff, 2019)
To determine the appropriateness of the grammatical and sentence structures of the M-PFPSLD questionnaire.	Comments and suggestions from the expert committees	Qualitative content validity analysis

RESULTS

This section covers presentation of the results for expert committee review (including qualitative and quantitative content validity), face validity and pretesting, as well as proofreading.

Stage 1: Expert Committee Review

For the expert committee review, the results were divided into three parts: i) overall review, ii) qualitative content validity analysis, and iii) quantitative content validity analysis.

i) Overall Review by Expert Committee

Overall, the translated version of the questionnaire (BM-H) and original version (E-H) could be considered satisfactory and comparable to the original version as there were only some minor amendments made to the BM-H version in terms of wording, grammar, sentence structure, and scoring that suit the culture in Malaysia (refer to Table 2). After amendments were made, BM-H was finalised into the second version of the Harmonised M-PFPSLD (BM-H2).

ii) Qualitative Content Validity Analysis

The qualitative content validity analysis was carried out by the expert committees and was measured through the recommendations given in terms of wording, grammar, sentence structure, and scoring (Safikhani et al., 2013). Therefore, there were some changes made to refine the structure and the content of the items. The changes made included 1) refining of phrases or sentences; 2) improvements in terms of grammatical aspects; 3) changing of word selections (please refer to Appendix C). There were 26 affected items out of 28 items that were minorly amended based on the discussion among the committee members (please refer to Appendix C) meanwhile only item 1 and item 9 were not included as there were no changes made. The examples of qualitative content validity analysis of several affected items in the BM-H version were included in Table 2. These improvements were made to achieve the equivalency between the original and translated version of the questionnaire before being finalised into the second version of the Harmonised M-PFPSLD (BM-H2) (please refer to Appendix D).

Table 2: The Examples of Qualitative Content Validity Analysis of BM-H Version

Item No.	Original Version	BM-H	BM-H2	Remarks on changes made
Item 2	I do not know how to react to some of the behaviors my child exhibits.	Saya tidak tahu **bagaimana untuk bertindak balas terhadap beberapa tingkah laku yang ditunjukkan oleh anak saya.**	Saya tidak tahu **cara untuk memberi reaksi kepada anak saya apabila dia menunjukkan tingkah-laku yang tidak patut dilakukan.**	• **Changed word selections** • **Refined sentences**
Item 3	I notice myself and my child becoming frustrated when we are unable to determine what the other is trying to say.	Saya perasan **diri saya** dan anak saya **menjadi kecewa** apabila kami tidak dapat **menentukan apa yang orang lain cuba sampaikan.**	Saya perasan **saya** dan anak saya **masing-masing berasa kecewa** apabila kami tidak dapat **memahami antara satu sama lain.**	• **Changed word selections** • **Refined phrases**
Item 4	My child's temper tantrums seem to occur more often and at a more severe level than typical developing children.	**Kemarahan** anak saya **sepertinya berlaku** lebih kerap dan **pada tahap yang lebih teruk berbanding kanak-kanak yang mempunyai perkembangan biasa.**	**Keadaan** anak saya **apabila mengamuk** lebih kerap dan **teruk daripada kanak-kanak yang mempunyai perkembangan normal.**	• **Changed word selections** • **Refined phrases**
Item 5	I feel as if other people expect to make decisions for my child rather than allowing them to make the decision on their own.	Saya rasa **seolah-olah** orang lain **mengharapkan saya untuk membuat keputusan** bagi pihak anak saya daripada membenarkan anak saya membuat keputusannya sendiri.	Saya rasa orang lain **cuba untuk membuat keputusan** bagi pihak anak saya daripada membenarkan anak saya membuat keputusannya sendiri.	• **Changed word selections** • **Refined phrases**

ii) Quantitative Content Validity Analysis

For this study, the I-CVI was calculated for all 28 items, and it was found that the I-CVI of each item was 80% and above (except for item 2 with the I-CVI of 0.5), and the average I-CVI (I-CVI/Ave) was 0.97. The I-CVI must be equal to or higher than 79% to be considered appropriate (Abdullah et al., 2010). This indicated that most of the items in the BM-H2 version were appropriate. For S-CVI, there were two methods to calculate it, which were S-CVI/Ave (sum of I-CVI scores divided by the number of items, which were 28) and S-CVI/UA (the items rated with a scoring of 3 or 4 by the expert committee divided by the total items, which were 28). The S-CVI/Ave and S-CVI/UA were found to be 0.93 and 0.71, respectively. A minimum score of 0.80 for both S-CVI/Ave and S-CVI/UA was considered acceptable, while a score of 0.90 and above would be considered excellent (Pilot et al., 2007, as cited in Haidari and Uzun, 2019). Therefore, the S-CVI/Ave showed an excellent congruity level among the content experts over the relevancy of all 28 items, while the S-CVI/UA was inadequate. This issue was considered to be solved after the revision and amendments were made in terms of wording selections and sentence structures. The detailed calculation of this I-CVI and S-CVI was shown in Table 3.

Stage 2: Face Validity and Pre-Testing

For this study, the researcher found that all the items in the BM-H2 version were acceptable with a 100% score on the face validity index (FVI) based on its criteria of readability, feasibility, clarity of words, and layout and style. The time taken to complete the initial version of M-PFPSLD was 5 to 30 minutes, as reported by all the participants. The average time taken to complete the BM-H2 version was 11.5 minutes. All participants reported that all the items were understandable. The M-PFPSLD questionnaire, after pre-testing, is attached in Appendix D.

Stage 3: Proofreading

Two Malay language experts proofread the BM-H2 version of the questionnaire. Both experts had been working as Bahasa Melayu teachers and had more than five years of teaching experience. After proofreading all the items in the initial version of the questionnaire, no changes were made, thus indicating that the BM-H2 version was well-written, grammatical-error-free, and accurately expressed the intended meaning of the original context. The finalised version of the M-PFPSLD questionnaire (BM-H2) could be found in Appendix D.

Table 3: The Summary of Quantitative Content Validity Analysis of BM-H2

Items No.	Expert 1	Expert 2	Expert3	Expert4	Expert5	Expert6	Experts in Agreement	I-CVI
1, 7-11, 13-17, 19-24, 26-28	√	√	√	√	√	√	6	1.00
3	√	√	√	√	√	**X**	5	0.83
4	√	√	√	√	√	**X**	5	0.83
5	√	**X**	√	√	√	√	5	0.83
6	√	√	√	**X**	√	√	5	0.83
12	√	**X**	√	√	√	√	5	0.83
18	√	√	√	√	√	**X**	5	0.83
25	√	**X**	√	√	√	√	5	0.83
2	√	**X**	√	√	**X**	**X**	3	0.50
I-CVI/Ave								0.96
S-CVI/Ave								0.93
S-CVI/UA								0.71
Proportion relevance	1.00	0.86	1.00	0.96	0.96	0.86		
Average proportion of items judged as relevant across the six experts							0.94	

*Note: √ = rating of 3 or 4; **X** = rating of 1 or 2; BM-H2 = Second Version of the Harmonized Malay Translation; I-CVI = Item-Level Content Validity Index; S-CVI = Scale-Level Content Validity Index; Ave = Average; UA = Universal Agreement*

DISCUSSION

This study was designed to validate the initial version of the M-PFPSLD (BM-H) questionnaire that may act as a preliminary psychometric property of the M-PFPSLD questionnaire. The translated and adapted version of the PFPSLD questionnaire (Jamaludin & Jusoh, 2023), BM-H was content-validated, face-validated, pre-tested, and proofread to certify the contents of the BM-H. This validation process produced the second version of the questionnaire (BM-H2) that may act as an initial step to prove the suitability of the translated and adapted contents to be used in the Malay-speaking population.

For this study, the procedure was done from the content validity analysis until proofreading of all 28 items of the questionnaire to produce the second harmonised version of the PFPSLD (BM-H2) questionnaire. Generally, the BM-H version appeared to be culturally acceptable, and all 28 items as per the translated and adapted version (Jamaludin & Jusoh, 2023) were maintained with minor amendments. The amendments were made to improve the sentence structures and to ensure the contents of the items in the questionnaire were comprehensible and culturally appropriate for the Malay-speaking population. This was done by refining phrases or sentences, improving grammatical aspects, and changing word selections. Given the difference in languages and cultures between the original sources and the target language in this study,

minor amendments were often necessary to maintain the conceptual and content equivalence between the original questionnaire and its translated version (Guillemin et al., 1993).

This study found that the content validity analysis to evaluate the inclusiveness and representativeness of the content of the M-PFPSLD questionnaire were good since the I-CVI values for each item and overall items were 0.83 to 1, which was higher than the recommended value of 0.79 to be considered as appropriate (Abdullah et al., 2010). However, the I-CVI value for item 2 was 0.50 due to the wording selections and sentence structure. This indicated that all items except for item 2 in the M-PFPSLD were appropriate in evaluating the parental feelings and perceptions towards their child's speech or language disorder. The amendments were made to item 2 by changing the words used and refining the sentences based on the comments from the content experts. Next, the scale-level content validity for the BM-H2 version was inadequate by using the S-CVI/UA approach (0.71) (Hadie et al., 2017; Ozair et al., 2017; Lau et al., 2018; Marzuki et al., 2018 cited in Yusoff, 2019), but it was adequate when using the S-CVI/Ave approach (0.93) (Pilot et al, 2007 cited in Haidari and Uzun, 2019). It was recommended that an S-CVI/Ave value of 0.93, higher than the minimum score of 0.80 or 0.90 (Pilot et al., 2007, cited in Haidari and Uzun, 2019), had achieved a satisfactory level of content validity. Therefore, the content validity results indicated that the M-PFPSLD had good comprehensiveness and was representative of the study population, which was the Malay-speaking population. No further revision was necessary for the items, as all of the items were reported to be clear and easy to comprehend.

The consistency of the CVI findings of this M-PFPSLD version could not be compared with other previous studies since there were limited studies that have conducted CVI analysis on the original version of the PFPSLD and the translated versions of the PFPSLD up to the researcher's knowledge. Thus, it was assumed that the CVI analysis on the PFPSLD scale had not been examined before. Besides, other than the PFPSLD scale, CVI analysis for other scales that measured the parents' feelings and perceptions, including the Parental Assessment Questionnaire for Children's Emotional Competence (URPEKD) (Milkovic et al., 2020) and the Parent and Child Forms of the Parent Perception Inventory (PPI-P and PPI-C) (Cole et al., 2018) were also not found. Milkovic et al. (2020) did mention content validation in their study; however, it was limited to the discussion on the content experts in agreement and no further discussion on the CVI analysis.

LIMITATIONS OF THE STUDY AND FUTURE RECOMMENDATIONS

There was a limitation in this study that should be acknowledged. This study was only conducted up to the proofreading process after face validity and pre-testing. This was because this study was continued from the content validation procedures, which influenced the time taken for the researcher, the expert committees, and the non-expert committees to complete the cross-cultural adaptation and validation process. Thus, the M-PFPSLD questionnaire was not yet ready to be used for the Malay-speaking population since the psychometric properties were not fully examined due to time constraints. Therefore, it was recommended for future studies to explore the psychometric properties in depth in terms of the pilot study and construct validity testing to produce the final version of the M-PFPSLD. This recommendation to further evaluate the psychometric properties will be beneficial to produce a valid and feasible M-PFPSLD questionnaire to be used in the Malay-speaking population.

CONCLUSION

In conclusion, this study provided evidence that the M-PFPSLD questionnaire had a good content validity index as one of the preliminary psychometric properties of the questionnaire. At the end of the study, the valid and feasible M-PFPSLD questionnaire will be produced to be used for Malay-speaking parents, as it will be relevant to the cultural and linguistic aspects of the Malay-speaking population. The findings might act as preliminary psychometric properties of the M-PFPSLD questionnaire, thus future studies can further research the psychometric properties of the M-PFPSLD, such as reliability, construct validity, convergent and divergent validity, and reliability testing. This study also acts as an initial step to gain insights into the parents' feelings and perceptions towards their child's speech or language disorder. This information might be useful for implementing and designing a treatment plan that better suits the children's speech or language disorders without overlooking the perspectives of their parents. Besides, the findings might be helpful for future research on speech-language areas in Malaysia. Therefore, this study would be beneficial for the professionals involved in the management of speech and language disorders, parents, as well as children with speech or language disorders, in reaching a suitable management that aligns with the child's needs.

ACKNOWLEDGEMENT

The authors would like to express their deepest gratitude to everyone who helped carry out this study successfully, either directly or indirectly. This research was not funded by any grant.

REFERENCES

Abidin R. R. (2012). Parenting Stress Index-Fourth Edition (PSI-4). *Psychological Assessment Resources, 33(7).* https://doi.org/10.1177/0734282914556069

American Speech-Language-Hearing Association [ASHA]. (1993). Definitions of Communication Disorders and Variations. *Definitions of Communication Disorders and Variations*. https://www.asha.org/policy/rp1993-00208/

Alias A. & Ramly U. (2021). Parental Involvement in Speech Activities of Speech Delayed Child at Home. *Advances in Social Science, Education and Humanities Research*, 540, 217-222.

Barrat, J., Khoza-Shangase, K., & Msimang, K. (2012). Speech-Language Assessment in A Linguistically Diverse Setting: Preliminary Exploration of the Possible Impact of Informal 'Solutions' Within the South African Context. *South African Journal of Communication Disorders, 59(1)*, 34-44. https://doi/org/10.7196/SAJCD.51

Belenkova, N., & Davtyan, V. (2016). Correlation of Translation and Other Language Activities. *International Journal of Environmental & Science Education, 11(18)*, 10951-10959. https://files.eric.ed.gov/fulltext/EJ1120579.pdf

Bell, M. L. et al. (2013). Differential Dropout and Bias in Randomised Controlled Trials: When It Matters and When It May Not. National Library of Medicine [NIH]. https://doi.org/10.1136/bmj.e8668

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioural Research: A Primer. *Frontiers in Public Health, 6,* 149. https://doi.org/10.3389/fpubh.2018.00149

Burn, R. B. (1995). "Introduction to Research Methods." Melbourne: Longman.

Cole, D. A. et al. (2018). Validating Parent and Child Forms of the Parent Perception Inventory. *Psychol Assess, 30(8),* 1065-1081. https://doi.org/10.1037/pas0000552

Donenberg, G., & Baker, B. L. (1993). The impact of young children with externalizing behaviors on their families. *Journal of Abnormal Child Psychology, 21*(2), 179–198. https://doi.org/10.1007/BF00911315

Foster, M., Choo, A., & Smith, S. (2023). Speech-Language Disorder Severity, Academic Success, and Socioemotional Functioning Among Multilingual and English Children in the United States: *The National Survey of Children's Health. Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1096145

Guillemin, F., Bombardier, C., & Beaton, D. E. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology, 46(12)*, 1417-1432. https://doi.org/10.1016/0895-4356(93)90142-n

Haidari, S. M. & Uzun, N. B. (2019). Content Validity and Confirmatory Factor Analysis of Cooperative Learning Attitude Scale (CLAS) for the EFL Students. *International Journal of Social Sciences and Education Research, 5(4),* 418-433. https://files.eric.ed.gov/fulltext/ED605793.pdf

Holley, E. (2018). Parents' Feelings and Perceptions Towards Their Child's Speech or Language Disorder. http://ideaexchange.uakron.edu/honors_research_projectshttp://ideaexchange.uakron.edu/honors_research_projects/609

Hurst, S. et al. (2015). Pretesting Qualitative Data Collection Procedures to Facilitate Methodological Adherence and Team Building in Nigeria. *International Journal of Qualitative Methods*, *14*, 53–64. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4393011/

Jamaludin, M., & Jusoh, M. (2023). Translation, Adaptation and Validation of Parents' Feelings and Perception Towards their Child's Speech or Language Disorder Questionnaire. IIUM FYP.

Jensen de López, K. M. et al. (2021). Exploring parental perspectives of childhood speech and language disorders across 10 countries: A pilot qualitative study. *Journal of Speech, Language, and Hearing Research*, *64*(5), 1739-1747. https://doi.org/10.1044/2020_JSLHR-20-00415

Kiogora, N. (2021). Children with Speech and Language Disorders. *International Academic Journal of Arts and Humanities (IAJAH), 1(2)*, 360-376. https://iajournals.org/articles/iajah_v1_i2_360_376.pdf

Krabbe, P. F. M. (2017). Chapter 7: Validity. *The Measurement of Health and Health Status*, 113-134. https://doi.org/10.1016/B978-0-12-801504-9.00007-6

Langbecker, D., Snoswell, C. L., Smith, A. C., Verboom, J., & Caffery, L. J. (2020). Long-Term Effects of Childhood Speech and Language Disorders: A Scoping Review. *South African Journal of Childhood Education, 10(1)*, 1-13. https://doi.org/10.4102/sajce.v10i1.801

Leonard, L. & Schroeder, M. (2023). The Study of Children with Developmental Language Disorder Beyond English: A Tutorial. *Language Acquisition*, 1-19. https://doi.org/10.1080/10489223.2023.2197891

Lyons, R. (2021). Impact of Language Disorders on Children's Everyday Lives from 4 to 13 Years: Commentary on Le, Mensah, Eadie, McKean, Schiberras, Bavin, Reilly and Gold. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 62(12)*, 1485-1487. https://doi.org/10.1111/jcpp.13391

McLeod, S. (2023, September 25). *Qualitative Vs Quantitative Research Methods & Data Analysis.* Simply Psychology. https://www.simplypsychology.org/qualitative-quantitative.html

Milkovic, A. N. et al. (2020). Construction and Validation of the Parental Assessment Questionnaire for Children's Emotional Competence (URPEKD). https://doi.org/10.32903/zs.66.1.1

Varni, J. W., Sherman, S., Burwinkle, T., Dickinson, P., & Dixon, P. (2004). The PedsQL 4.0 Family Impact Module: Reliability and Validity. *Journal of Pediatric Psychology, 36(5),* 517-528. https://doi.org/10.1186/1477-7525-2-55

Silva, L., Lamonica, D., & Hage, S. (2021). Instruments to Assess Children's Language and Speech Translated and Adapted into Brazilian Portuguese: *An Integrative. Revista CEFAC, 23(2)*. https://doi.org/10.1590/1982-0216/202123212520

Simkus, J. (2023, July 31). *Convenience sampling: Definition, method and examples*. Simply Psychology. https://www.simplypsychology.org/convenience-sampling.html

Statista Research Department. (2023, September 13). *Malaysia: breakdown of population by ethnicity 2019-2020*. Statista. https://www.statista.com/statistics/1017372/malaysia-breakdown-of-population-by-ethnicity

Sun, J. & Lee, E. (2022). Parents' and Teachers' Perceptions of Speech-Language Rehabilitation and Demands for the Use of Treatment-Related Apps. *Clinical Archives of Communication Disorders, 7(3)*, 137-143. https://doi.org/10.21849/cacd.2022.00892

Tee-Melegrito, R. A. (2022, April 27). *Communication disorders: Definition, types, causes, and symptoms*. Medicalnewstoday.com. https://www.medicalnewstoday.com/articles/communication-disorders

Uysal, A. A., & Tura, G. (2019). The Views and Knowledge of Parents of Children with Speech/ Language Disorders on Speech and Language Therapy in Turkey. *Examines in Physical Medicines and Rehabilitation, 2(2)*, 1-4. https://doi.org/10.31031/EPMR.2019.02.000533

Wang, X., & Cheng, Z. (2020). Cross-Sectional Studies: Strengths, Weaknesses, and Recommendations. *Chest*, *158(1)*, 65–71. https://doi.org/10.1016/j.chest.2020.03.012

Yusoff M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal, 11(2)*, 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Yusoff M. S. B. (2019). ABC of response process validation and face validity index calculation. *Education in Medicine Journal, 11(3)*, 55–61. https://doi.org/10.21315/eimj2019.11.3.6

APPENDIX A: ORIGINAL VERSION OF PFPSLD QUESTIONNAIRE

Background Information:
Child Age: _____
Diagnosed Disorder of Child: _____
How long has your child had a speech or language disorder? _____

Please circle the most appropriate answer in regards to the questions below:

1. I worry about my child's future.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

2. I do not know how to react to some of the behaviors my child exhibits.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

3. I notice myself and my child becoming frustrated when we are unable to determine what the other is trying to say.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

4. My child's temper tantrums seem to occur more often and at a more severe level than typical developing children.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

5. I feel as if other people expect to make decisions for my child rather than allowing them to make the decision on their own.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

6. I feel as if my relationship with my child is affected because of their communication difficulties**.**
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

7. I would be open to allowing my child to use a form of augmentative and alternative Communication (AAC) as a form of support.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

Please explain why or why not:
_____
_____

8. I feel the need to help my child with certain things more than I should.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

9. I often find myself wanting to speak for my child if she or he is having trouble communicating with someone.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

10. I believe I could have done something differently in my child's early development to prevent him or her from obtaining a speech or language disorder.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

11. I worry sometimes that my child will never improve.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

12. I think that my child's communication disorder could have been affected by his or personality at a young age.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

13. I feel as if I could have provided a more stimulating language environment for my child that could have prevented his or her communication disorder.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

14. I do not think my child will have the same achievements as his or her typically developing peers.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

15. I think my child's communication disorder causes them to have more severe behavior issues than typically developing children.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

16. I feel as if my child is timid to speak up because of his or her communication disorder.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

17. I feel as if my child has less independence because of his or her communication disorder.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

18. I feel as if my child is at a disadvantage for developing social relationship with other children of his or her age.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

19. I would describe my child as a loner, or someone who distances themselves from others.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

20. I feel as if my child's communication disorder causes him or her to be singled out by others.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

21. I fear that my child could be bullied by other children because they are seen as "different."
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

22. I feel as if my child does not receive the educational support needed for his or her speech or language disorder.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

23. I am afraid my child could be placed in an educational setting that is not suitable to his or her needs in regards to their communication disorder
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

24. I feel as if my child is having a harder time in school as his or her typical developing peers.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

25. I feel as if my child needs to be in a stable environment to ensure success in his or her education.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

26. I feel as if my child has a hard time paying attention in class due to the distractions surrounding them.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

27. I feel as if my child has a hard time expressing his or her feelings to others.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

28. I worry that my child will be in a harmful situation and be unable to communicate with anyone.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

APPENDIX B: INITIAL VERSION OF M-PFPSLD QUESTIONNAIRE

Harmonised version of Malay-PFPSLD (BM-H PFPSLD)

Parents Feelings and Perceptions Towards their Child's Speech or Language Disorder (PFPSLD) Questionnaire - ***Soal Selidik Perasaan dan Persepsi Ibu Bapa Terhadap Kecelaruan Pertuturan atau Bahasa Anak (M-PFPSLD)***

Harmonised Malay version
Maklumat latar belakang
Umur anak
Diagnosis kecelaruan pertuturan atau bahasa anak
Sudah berapa lamakah anak anda mempunyai kecelaruan pertuturan atau bahasa?
Arahan: Sila bulatkan jawapan yang paling sesuai berkaitan soalan soalan di bawah.
Pemarkahan: Sangat Setuju Setuju Neutral Tidak Setuju Sangat Tidak Setuju
1. Saya bimbang tentang masa depan anak saya.
2. Saya tidak tahu bagaimana untuk bertindak balas terhadap beberapa tingkah laku yang ditunjukkan oleh anak saya.
3. Saya perasan diri saya dan anak saya menjadi kecewa apabila kami tidak dapat menentukan apa yang orang lain cuba sampaikan.
4. Kemarahan anak saya sepertinya berlaku lebih kerap dan pada tahap yang lebih teruk berbanding kanak-kanak yang mempunyai perkembangan biasa.
5. Saya rasa seolah-olah orang lain mengharapkan saya untuk membuat keputusan bagi pihak anak saya daripada membenarkan anak saya membuat keputusannya sendiri.
6. Saya rasa seolah-olah hubungan saya dengan anak saya terjejas disebabkan masalah komunikasi anak saya.
7. Saya terbuka untuk membenarkan anak saya menggunakan komunikasi tambahan dan alternatif (*Augmentative and Alternative Communication*, AAC) sebagai satu bentuk sokongan. Sila jelaskan mengapa atau mengapa tidak.
8. Saya merasakan perlu untuk membantu anak saya dengan perkara-perkara tertentu lebih daripada yang sepatutnya saya lakukan.
9. Saya sering mendapati diri saya ingin bercakap bagi pihak anak saya jika dia mengalami kesukaran berkomunikasi dengan seseorang.

10. Saya percaya saya boleh melakukan sesuatu yang berbeza dalam perkembangan awal anak saya untuk mencegahnya daripada mendapat kecelaruan pertuturan atau bahasa.
11. Kadang-kadang, saya bimbang anak saya tidak akan menunjukkan perkembangan yang baik.
12. Saya berpendapat bahawa kecelaruan komunikasi anak saya mungkin telah terjejas disebabkan oleh personalitinya pada waktu kecil dahulu.
13. Saya rasa seolah-olah saya boleh menyediakan persekitaran bahasa yang lebih memberangsangkan bagi mencegah kecelaruan komunikasinya.
14. Saya tidak fikir anak saya akan mempunyai pencapaian yang sama seperti rakan sebayanya yang mempunyai perkembangan biasa.
15. Saya berpendapat bahawa kecelaruan komunikasi anak saya telah menyebabkan mereka mengalami masalah tingkah laku yang lebih teruk berbanding kanak-kanak lain yang mempunyai perkembangan biasa.
16. Saya rasa seolah-olah anak saya malu untuk bersuara kerana kecelaruan komunikasinya.
17. Saya rasa seolah-olah anak saya kurang berdikari kerana kecelaruan komunikasinya.
18. Saya rasa seolah-olah anak saya kurang bernasib baik untuk membina hubungan sosial dengan kanak-kanak lain yang sebaya dengannya.
19. Saya menggambarkan anak saya sebagai seorang yang menyendiri, atau seseorang yang menjauhkan dirinya daripada orang lain.
20. Saya rasa seolah-olah kecelaruan komunikasi anak saya menyebabkan dia dipulaukan oleh orang lain.
21. Saya takut anak saya akan dibuli oleh kanak-kanak lain kerana mereka dilihat sebagai "berbeza".
22. Saya rasa seolah-olah anak saya tidak mendapat sokongan pendidikan yang diperlukan untuk kecelaruan pertuturan atau bahasanya.
23. Saya takut anak saya akan ditempatkan dalam suasana pendidikan yang tidak sesuai dengan keperluan kecelaruan komunikasinya.
24. Saya rasa seolah-olah anak saya mengalami masa yang lebih sukar di sekolah berbanding rakan sebayanya yang mempunyai.
25. Saya rasa seolah-olah anak saya perlu berada dalam persekitaran yang stabil untuk memastikan kejayaan dalam pendidikannya.
26. Saya rasa seolah-olah anak saya mengalami kesukaran untuk menumpukan perhatian di dalam kelas disebabkan oleh gangguan di sekeliling mereka.
27. Saya rasa seolah-olah anak saya mengalami kesukaran untuk meluahkan perasaannya kepada orang lain.
28. Saya bimbang anak saya akan berada dalam situasi yang berbahaya dan tidak dapat berkomunikasi dengan sesiapa pun.

APPENDIX C: QUALITATIVE CONTENT VALIDITY ANALYSIS OF BM-H

Item No.	Original Version	BM-H	BM-H2	Remarks on Changes Made
Item 1	I worry about my child's future.	Saya bimbang tentang masa depan anak saya.	Saya bimbang tentang masa depan anak saya.	**No changes**
Item 2	I do not know how to react to some of the behaviors my child exhibits.	Saya tidak tahu **bagaimana** untuk **bertindak balas terhadap beberapa tingkah laku yang ditunjukkan oleh anak saya.**	Saya tidak tahu **cara** untuk **memberi reaksi kepada anak saya apabila dia menunjukkan tingkah-laku yang tidak patut dilakukan.**	● **Changed word selections** ● **Refined sentences**
Item 3	I notice myself and my child becoming frustrated when we are unable to determine what the other is trying to say.	Saya perasan **diri saya** dan anak saya **menjadi** kecewa apabila kami tidak dapat **menentukan apa yang orang lain cuba sampaikan.**	Saya perasan **saya** dan anak saya **masing-masing berasa** kecewa apabila kami tidak dapat **memahami antara satu sama lain.**	● **Changed word selections** ● **Refined phrases**
Item 4	My child's temper tantrums seem to occur more often and at a more severe level than typical developing children.	**Kemarahan anak saya sepertinya berlaku lebih kerap dan pada tahap yang lebih teruk berbanding** kanak-kanak yang mempunyai perkembangan **biasa**.	**Keadaan anak saya apabila mengamuk lebih kerap dan teruk daripada** kanak-kanak yang mempunyai perkembangan **normal**.	● **Changed word selections** ● **Refined phrases**
Item 5	I feel as if other people expect to make decisions for my child rather than allowing them to make the decision on their own.	Saya rasa **seolah-olah** orang lain **mengharapkan saya** untuk membuat keputusan bagi pihak anak saya daripada membenarkan anak saya membuat keputusannya sendiri.	Saya rasa orang lain **cuba** untuk membuat keputusan bagi pihak anak saya daripada membenarkan anak saya membuat keputusannya sendiri.	● **Changed word selections** ● **Refined phrases**
Item 6	I feel as if my relationship with my child is affected because of their communication difficulties.	Saya rasa **seolah-olah** hubungan saya dengan anak saya terjejas **disebabkan masalah komunikasi anak saya.**	Saya rasa hubungan saya dan anak saya terjejas **disebabkan oleh kesukarannya untuk berkomunikasi.**	● **Refined sentences** ● **Improved grammar**
Item 7	I would be open to allowing my child to use a form of augmentative and alternative Communication (AAC) as a form of support. Please explain why or why not:	Saya terbuka untuk membenarkan anak saya menggunakan komunikasi tambahan dan alternatif (*Augmentative and Alternative Communication*, AAC) sebagai satu bentuk sokongan. Sila jelaskan **mengapa atau mengapa tidak.**	Saya terbuka untuk membenarkan anak saya menggunakan komunikasi tambahan dan alternatif (*Augmentative and Alternative Communication*, AAC) sebagai satu bentuk sokongan. Sila jelaskan **sebab anda menyatakan demikian.**	**Refined sentences**
Item 8	I feel the need to help my child with certain things more than I should.	Saya rasa **perlu** untuk membantu anak saya **dengan perkara-perkara tertentu** lebih daripada yang sepatutnya saya lakukan.	Saya rasa **keperluan** untuk membantu anak saya **melakukan sesuatu perkara** lebih daripada yang sepatutnya saya lakukan.	● **Improved grammar** ● **Refined sentences**

Item 9	I often find myself wanting to speak for my child if she or he is having trouble communicating with someone.	Saya sering mendapati diri saya ingin bercakap bagi pihak anak saya jika dia mengalami kesukaran berkomunikasi dengan seseorang.	Saya sering mendapati diri saya ingin bercakap bagi pihak anak saya jika dia mengalami kesukaran berkomunikasi dengan seseorang.	**No changes**
Item 10	I believe I could have done something differently in my child's early development to prevent him or her from obtaining a speech or language disorder.	Saya percaya **saya boleh melakukan sesuatu yang berbeza dalam** perkembangan awal anak saya untuk **mencegahnya** daripada mendapat kecelaruan pertuturan atau bahasa.	Saya percaya **bahawa saya sepatutnya boleh melakukan sesuatu yang berbeza semasa** perkembangan awal anak saya untuk **menghindarinya** daripada mendapat kecelaruan pertuturan atau bahasa.	● **Refined sentences** ● **Changed word selections**
Item 11	I worry sometimes that my child will never improve.	**Kadang-kadang**, saya **bimbang anak saya** tidak akan menunjukkan **perkembangan yang baik.**	**Adakalanya** saya **berasa bimbang jika anak saya** tidak akan menunjukkan **perubahan untuk pulih.**	● **Changed word selections** ● **Refined phrases**
Item 12	I think that my child's communication disorder could have been affected by his or personality at a young age.	Saya berpendapat bahawa kecelaruan komunikasi anak saya **mungkin telah terjejas** disebabkan oleh personalitinya **pada waktu kecil dahulu.**	Saya berpendapat bahawa kecelaruan komunikasi anak saya **terjejas** disebabkan oleh personalitinya **sewaktu dia kecil.**	**Refined sentences**
Item 13	I feel as if I could have provided a more stimulating language environment for my child that could have prevented his or her communication disorder.	Saya rasa **seolah-olah** saya boleh menyediakan persekitaran bahasa yang lebih memberangsangkan **bagi mencegah kecelaruan komunikasinya.**	Saya rasa saya boleh menyediakan persekitaran bahasa yang lebih memberangsangkan **untuk anak saya bagi menghindari kecelaruan komunikasinya.**	● **Improved grammar** ● **Changed word selections** ● **Refined sentences**
Item 14	I do not think my child will have the same achievements as his or her typically developing peers.	Saya tidak **fikir** anak saya akan **mempunyai** pencapaian yang sama seperti rakan sebayanya yang mempunyai perkembangan **biasa**.	Saya tidak **yakin** anak saya akan **memperoleh** pencapaian yang sama seperti rakan sebayanya yang mempunyai perkembangan **yang normal.**	**Changed word selections**
Item 15	I think my child's communication disorder causes them to have more severe behavior issues than typically developing children.	Saya berpendapat bahawa kecelaruan komunikasi **anak saya** telah menyebabkan **mereka mengalami masalah** tingkah laku yang lebih teruk berbanding **kanak-kanak lain** yang mempunyai perkembangan **biasa**.	Saya berpendapat bahawa kecelaruan komunikasi **yang dialami oleh anak saya** telah menyebabkan **dia mempunyai masalah** tingkah laku yang lebih teruk berbanding r**akan sebayanya** yang mempunyai perkembangan **normal.**	● **Changed word selections** ● **Refined phrases**
Item 16	I feel as if my child is timid to speak up because of his or her communication disorder.	Saya rasa **seolah-olah** anak saya **malu** untuk **bersuara kerana kecelaruan komunikasinya.**	Saya rasa anak saya **takut** untuk **memberi pandangan disebabkan oleh kecelaruan komunikasi yang dialaminya.**	● **Improved grammar** ● **Changed word selections** ● **Refined sentences**

Item 17	I feel as if my child has less independence because of his or her communication disorder.	Saya rasa **seolah-olah** anak saya **kurang berdikari kerana kecelaruan komunikasinya.**	Saya rasa anak saya **mempunyai daya berdikari yang rendah disebabkan oleh kecelaruan komunikasi yang dialaminya.**	● **Improved grammar** ● **Refined sentences**
Item 18	I feel as if my child is at a disadvantage for developing social relationship with other children of his or her age.	Saya rasa **seolah-olah** anak saya **kurang bernasib baik** untuk membina hubungan sosial dengan kanak-kanak lain yang **sebaya dengannya.**	Saya rasa anak saya **kurang berkebolehan** untuk membina hubungan sosial dengan kanak-kanak lain yang **seusia dengannya.**	● **Improved grammar** ● **Refined phrases**
Item 19	I would describe my child as a loner, or someone who distances themselves from others.	Saya **menggambarkan** anak saya sebagai **seorang** yang menyendiri, atau seseorang yang menjauhkan dirinya daripada orang lain.	Saya **gambarkan** anak saya sebagai **seseorang** yang suka menyendiri, atau seseorang yang suka menjauhkan dirinya daripada orang lain.	● **Improved grammar** ● **Changed word selections**
Item 20	I feel as if my child's communication disorder causes him or her to be singled out by others.	Saya rasa **seolah-olah** kecelaruan komunikasi anak saya menyebabkan dia dipulaukan oleh orang lain.	Saya rasa kecelaruan komunikasi anak saya menyebabkan dia dipulaukan oleh orang lain.	**Improved grammar**
Item 21	I fear that my child could be bullied by other children because they are seen as "different."	Saya takut anak saya akan dibuli oleh kanak-kanak lain kerana mereka **dilihat sebagai "berbeza".**	Saya takut anak saya akan dibuli oleh kanak-kanak lain kerana mereka **melihatnya sebagai seseorang yang "berlainan".**	● **Refined phrases** ● **Changed word selections**
Item 22	I feel as if my child does not receive the educational support needed for his or her speech or language disorder.	Saya rasa **seolah-olah** anak saya tidak mendapat sokongan pendidikan yang diperlukan untuk kecelaruan pertuturan atau **bahasanya**.	Saya rasa anak saya tidak mendapat sokongan pendidikan yang diperlukan untuk kecelaruan pertuturan atau **bahasa yang dialaminya**.	● **Improved grammar** ● **Refined phrases**
Item 23	I am afraid my child could be placed in an educational setting that is not suitable to his or her needs in regards to their communication disorder.	Saya takut anak saya akan ditempatkan dalam **suasana** pendidikan yang tidak sesuai dengan **keperluan kecelaruan komunikasinya.**	Saya takut anak saya akan ditempatkan dalam **penempatan** pendidikan yang tidak sesuai dengan **keperluannya yang mengalami kecelaruan komunikasi.**	● **Chnaged word selections** ● **Refined sentences**
Item 24	I feel as if my child is having a harder time in school as his or her typical developing peers.	Saya rasa **seolah-olah** anak saya mengalami **masa yang lebih sukar** di sekolah berbanding rakan sebayanya yang mempunyai perkembangan **biasa**.	Saya rasa anak saya mengalami **kesukaran** di sekolah berbanding rakan sebayanya yang mempunyai perkembangan **normal**.	● **Improved grammar** ● **Changed word selections**
Item 25	I feel as if my child needs to be in a stable environment to ensure success in his or her education.	Saya rasa **seolah-olah** anak saya perlu berada dalam persekitaran yang stabil untuk memastikan kejayaan dalam pendidikannya.	Saya rasa anak saya perlu berada dalam persekitaran yang stabil untuk memastikan kejayaan dalam pendidikannya.	**Improved grammar**

Item 26	I feel as if my child has a hard time paying attention in class due to the distractions surrounding them.	Saya rasa **seolah-olah** anak saya mengalami kesukaran untuk menumpukan perhatian di dalam kelas disebabkan oleh **gangguan di sekeliling mereka.**	Saya rasa anak saya mengalami kesukaran untuk menumpukan perhatian di dalam kelas disebabkan oleh **persekitarannya yang mengganggu.**	• **Improved grammar** • **Refined sentences**
Item 27	I feel as if my child has a hard time expressing his or her feelings to others.	Saya rasa **seolah-olah** anak saya mengalami kesukaran untuk meluahkan perasaannya kepada orang lain.	Saya rasa anak saya mengalami kesukaran untuk meluahkan perasaannya kepada orang lain.	**Improved grammar**
Item 28	I worry that my child will be in a harmful situation and be unable to communicate with anyone.	Saya bimbang **anak saya** akan berada dalam situasi yang berbahaya **dan** tidak dapat berkomunikasi dengan sesiapa pun.	Saya bimbang **jika anak saya** akan berada dalam situasi yang berbahaya **namun** tidak dapat berkomunikasi dengan sesiapa pun.	**Changed word selections**

APPENDIX D: SECOND VERSION OF THE HARMONISED TRANSLATION (BM-H2)

Soal Selidik Perasaan dan Persepsi Ibu Bapa Terhadap Kecelaruan Pertuturan atau Bahasa Anak (M-PFPSLD)

Maklumat latar belakang:
Umur anak: _____
Diagnosis kecelaruan pertuturan atau bahasa anak: _____
Sudah berapa lamakah anak anda mempunyai kecelaruan pertuturan atau bahasa? _____

Arahan:
Sila bulatkan jawapan yang paling sesuai berkaitan soalan-soalan di bawah.

No.	Item	Pemarkahan				
		Sangat Setuju	**Setuju**	**Neutral**	**Tidak Setuju**	**Sangat Tidak Setuju**
1.	Saya bimbang tentang masa depan anak saya.					
2.	Saya tidak tahu cara untuk memberi reaksi kepada anak saya apabila dia menunjukkan tingkah-laku yang tidak patut dilakukan.					
3.	Saya perasan saya dan anak saya masing-masing berasa kecewa apabila kami tidak dapat memahami antara satu sama lain.					
4.	Keadaan anak saya apabila mengamuk lebih kerap dan teruk daripada kanak-kanak yang mempunyai perkembangan normal.					
5.	Saya rasa orang lain cuba untuk membuat keputusan bagi pihak anak saya daripada membenarkan anak saya membuat keputusannya sendiri.					
6.	Saya rasa hubungan saya dan anak saya terjejas disebabkan oleh kesukarannya untuk berkomunikasi.					
7.	Saya terbuka untuk membenarkan anak saya menggunakan komunikasi tambahan dan alternatif (*Augmentative and Alternative Communication*, AAC) sebagai satu bentuk sokongan. Sila jelaskan sebab anda menyatakan demikian: _____ _____ _____					
8.	Saya rasa keperluan untuk membantu anak saya melakukan sesuatu perkara lebih daripada yang sepatutnya saya lakukan.					

9.	Saya sering mendapati diri saya ingin bercakap bagi pihak anak saya jika dia mengalami kesukaran berkomunikasi dengan seseorang.					
10.	Saya percaya bahawa saya sepatutnya boleh melakukan sesuatu yang berbeza semasa perkembangan awal anak saya untuk menghindarinya daripada mendapat kecelaruan pertuturan atau bahasa.					
11.	Adakalanya saya berasa bimbang jika anak saya tidak akan menunjukkan perubahan untuk pulih.					
12.	Saya berpendapat bahawa kecelaruan komunikasi anak saya terjejas disebabkan oleh personalitinya sewaktu dia kecil.					
13.	Saya rasa saya boleh menyediakan persekitaran bahasa yang lebih memberangsangkan untuk anak saya bagi menghindari kecelaruan komunikasinya.					
14.	Saya tidak yakin anak saya akan memperoleh pencapaian yang sama seperti rakan sebayanya yang mempunyai perkembangan yang normal.					
15.	Saya berpendapat bahawa kecelaruan komunikasi yang dialami oleh anak saya telah menyebabkan dia mempunyai masalah tingkah laku yang lebih teruk berbanding rakan sebayanya yang mempunyai perkembangan normal					
16.	Saya rasa anak saya takut untuk memberi pandangan disebabkan oleh kecelaruan komunikasi yang dialaminya.					
17.	Saya rasa anak saya mempunyai daya berdikari yang rendah disebabkan oleh kecelaruan komunikasi yang dialaminya.					
18.	Saya rasa anak saya kurang berkebolehan untuk membina hubungan sosial dengan kanak-kanak lain yang seusia dengannya.					
19.	Saya gambarkan anak saya sebagai seseorang yang suka menyendiri, atau seseorang yang suka menjauhkan dirinya daripada orang lain.					
20.	Saya rasa kecelaruan komunikasi anak saya menyebabkan dia dipulaukan oleh orang lain.					
21.	Saya takut anak saya akan dibuli oleh kanak-kanak lain kerana mereka melihatnya sebagai seseorang yang "berlainan”.					
22.	Saya rasa anak saya tidak mendapat sokongan					

	pendidikan yang diperlukan untuk kecelaruan pertuturan atau bahasa yang dialaminya.					
23.	Saya takut anak saya akan ditempatkan dalam penempatan pendidikan yang tidak sesuai dengan keperluannya yang mengalami kecelaruan komunikasi.					
24.	Saya rasa anak saya mengalami kesukaran di sekolah berbanding rakan sebayanya yang mempunyai perkembangan normal.					
25.	Saya rasa anak saya perlu berada dalam persekitaran yang stabil untuk memastikan kejayaan dalam pendidikannya.					
26.	Saya rasa anak saya mengalami kesukaran untuk menumpukan perhatian di dalam kelas disebabkan oleh persekitarannya yang mengganggu.					
27.	Saya rasa anak saya mengalami kesukaran untuk meluahkan perasaannya kepada orang lain.					
28.	Saya bimbang jika anak saya akan berada dalam situasi yang berbahaya namun tidak dapat berkomunikasi dengan sesiapa pun.					

ORIGINAL ARTICLE

Psychometric Properties of The Malay Version-Revised Five-To-Fifteen (5-15R) Parental Questionnaire (Malay-5-15R)

Nurul Najwa Md Yusof[1], Masnira Jusoh[2,*]

[1]Speecheese Talent & Speech Centre, Level 1, Lot 18-A, Jalan Diplomatik 3, Presint 15, 62050 Putrajaya, W.P. Putrajaya

[2]Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
five-to-fifteen; 5-15R questionnaire; pre-testing; face validity; pilot study

ABSTRACT

Background: There are limited available or published questionnaires that suit the Malay-speaking population to screen or assess child development, such as physical development, behaviour, emotions, cognition, communication skills, or socialisation. Standardised questionnaires completed by parents and teachers can provide different perspectives on the children with behaviour and developmental needs. It can be a supporting assessment for the professionals to support the diagnosis of the children. Thus, the current study focuses on other psychometric properties to validate the initial version of Malay-5-15R. **Methods:** The initial version of the Malay Version-Revised Five-To-Fifteen (5-15R) questionnaire has been content validated by the experts from the previous study. The validity of Malay-5-15R was further examined by conducting 1) pre-testing and face validity; 2) proofreading; and 3) pilot study. **Results:** The initial version of the Malay-5-15R questionnaire was found to be easily understood by participants during pre-testing. A pilot study was conducted on the two groups of population: parents with normally developing children and parents with children having a secondary diagnosis. Results showed that the initial version of the Malay-5-15R exhibited sufficient internal consistency ($\alpha > 0.93$) and demonstrated an excellent coefficient of stability with a correlation coefficient value of 0.98. For convergent validity, most of the Malay-5-15R items have a significant correlation with items in the sub-domains, with r-values of greater than 0.30 ($p < 0.05$). Regarding divergent validity, about 99% of the items were correlated with r-values below 0.85 ($p < 0.05$). **Conclusion:** This study concluded that the Malay 5-15R questionnaire has good face validity, internal consistency reliability, test-retest reliability, convergent validity, and divergent validity, which are part of the psychometric properties. Therefore, it shows that the Malay 5-15R questionnaire is valid and reliable for use in the Malay-speaking population.

INTRODUCTION

Parents nowadays are concerned about their child's development. Early developmental milestones are the most significant period to be tracked to promote normal, typical development in children. The probability of upcoming health issues and educational challenges rises drastically if one or more developmental areas are delayed, such as physical development, behaviour, emotions, cognition, communication skills, or socialisation (Camden et al., 2019). According to Khan and Leventhal (2020), developmental screening needs to occur at the onset of clinical concerns, in response to parental worries, during the critical age between 0;09 to 2;05 years old.

For example, autism-specific screening tests are recommended at 18 and 24 months. A thorough assessment and follow-up are necessary because variations in the environment can cause developmental delays or they can be a sign of a developmental problem such as cerebral palsy, autism, or sensory or cognitive impairment. Therefore, the use of validated screening tools for children can help in the earlier identification of significant developmental delays. According to Shaw et al. (2023), it is critical to recognise neurodevelopmental disorders at an early age to guarantee that children have access to any services to help them develop those abilities.

Early identification offers the advantage of early intervention, which improves adaptive abilities and developmental outcomes. Early detection enables the utilisation of specialised interventions that address social communication, language development, and behavioural issues and are catered to the distinctive needs of individuals with autism spectrum disorder. Additionally, early diagnosis facilitates improved coping techniques, lowers anxiety among parents, and increases adult independence by giving families access to the right support

* *Corresponding author.*
E-mail address: masnirajusoh@iium.edu.my

services, educational materials, and community programmes (Okoye et al., 2023; Shaw et al., 2023).

LITERATURE REVIEW

Using structured, standardised parent or teachers' questionnaires has become more prominent in the early identification of developmental delays. These tools are affordable in terms of both professional resources and time, and they are straightforward to administer. Evaluations by parents are reasonably accurate when it comes to behavioural issues and symptoms (Korkman et al., 2004). When assessing children and adolescents with multiple mental health disorders, behavioural and learning difficulties, questionnaires and rating scales with strong psychometric properties have proven to be useful supplements to testing and interviews (Bohlin & Janols, 2004). Questionnaires can serve as a tool for health professionals to evaluate and as a screening tool for parents before deciding to undergo any formal evaluation by related professionals. These can be supplementary findings for health professionals rather than just assessing the child in the clinical setting. In Malaysia, there is a limited number of available or published questionnaires that suit the Malay-speaking population.

The Five to Fifteen (FTF) Questionnaire is one of the parental questionnaires that provides insights into children's behaviour and developmental problems. It was developed since 2004 by Björn Kadesjö, Lars-Olof Janols and Christopher Gillberg (Sweden); Marit Korkman and Katarina Mickelsson (Finland); Gerd Strand (Norway) and Anegen Trillingsgaard (Denmark). The purpose of the questionnaire is to lead the professionals to identify and measure cognitive, language, and motor impairment as well as social, emotional and behavioural problems of children and adolescents (Kadesjö et al., 2004).

The main objective of the FTF Questionnaire is to help parents evaluate how well their children function on every occasion, including their abilities and weaknesses. The parents can make developmental comparisons between their children and typical peers of their age before further assessment. It is shown that parents are reliable and valid sources of providing information about their child's development. Their early detection can contribute to early intervention and better prognosis (González et al., 2021). Other questionnaires have a similar purpose as the FTF Questionnaire, such as the Child Behaviour Analysis (CBCL), Child and Adolescent Psychiatric Assessment (CAPA), Devereux Scales of Mental Disorders (DSMD), and Behavioural Assessment System for Children (BASC) (Angold et al., 1995; Naglieri et al., 1994; Reynolds & Kamphaus, 1992; Achenbach, 1991).

Currently, the FTF Questionnaire has been revised since 2017 by the Nordic research group to enhance the comprehensiveness of the parents' questionnaire. The term "FTF" is also being substituted into "5-15R". The revised questionnaire includes the teachers' questionnaire, as the child behaviours can also be evaluated outside the family context. The questionnaire also extended the age range from five to fifteen years old and is currently applicable to children aged five to seventeen years old, as children's developmental issues need continuous monitoring from time to time. Impact questions were also included, as there is inconsistent information regarding the symptoms with the parents' experience of impairment (Kadesjö et al., 2017).

The 5-15R questionnaire was chosen in this study because it captures parents' perceptions of the abilities and shortcomings of their children across a broader range of developmental areas. Specifically, it encompasses eight major domains: motor skills, executive function, perception, memory, language, learning, social skills, and emotional/behavioural problems (Kadesjö et al., 2017). Other questionnaires focus solely on behaviour/emotion, such as CBCL, DSMD, and BASC. While Five-to-Fifteen-Revised (5-15R) ((previously known as FTF Questionnaire) covers a range of developmental domains. While the CAPA is very detailed for psychiatric diagnoses and the DSMD for severe disorders, the 5-15R offers a more general developmental assessment, making it more versatile for various issues. 5-15R provides detailed scoring and feedback for targeted interventions, offering more specific feedback in certain developmental areas compared to other tools.

The domains in the 5-15R Questionnaire are suitable to evaluate children with neurodevelopmental or neuropsychiatric disorder especially for attention deficit hyperactivity disorder, autism spectrum disorders (ASD), tic disorders, developmental coordination disorder (DCD), oppositional defiant disorder (ODD), conduct disorder (CD), emotional disorder, learning problems, or symptoms in these areas (Lindblad et al., 2011).

The 5-15R questionnaire has been translated and adapted from English to the Malay language in the previous study (Zainal & Jusoh, 2023). However, the adaptation process is only completed until content validation. To use the Malay-5-15R as an evaluation or screening tool in the Malay-speaking population, further psychometric features must be established, as the validation procedure is still progressing. Other than that, there was a limited number of available or published parental questionnaires that have

been translated into Malay, especially related to the developmental and behavioural issues of the children.

METHODOLOGY

The initial version of Malay-5-15R has been translated, adapted, and validated for content by Zainal and Jusoh (2023). Thus, the current study aims to continue the validation process (i.e., face validity and pre-testing processes) to improve the questionnaire's validity and identify problems that might only be noticeable to the targeted population during answering the questionnaire (Reynolds et al., 1993). Next, the proofreading process is needed to improve the sentence structure, grammar, and wording of the questionnaire by Bahasa Melayu experts among teachers and lecturers. Then, pilot testing will be conducted in two periods of time, with a two-week gap. The pilot study mainly examines the internal consistency reliability, test-retest reliability and inter-item correlation (to assess the convergent and divergent validity) within and between domains in Malay-5-15R.

Sample Size

The subject of the research study was recruited through purposive sampling. It is a non-probability sampling technique that aims to help answer research questions by concentrating on specific aspects of a population that are of interest (Rai & Thapa, 2015). The subjects were chosen based on the inclusion criteria.

According to Reynolds et al. (1993) and Julious (2005), a minimum sample size of 5 to 10 subjects is suggested for pre-testing and face validity, particularly for healthcare-related surveys. In a pilot study, it is suggested that a minimum sample size of 12 subjects is needed so that it is ready to be used in the Malaysian population (In, 2017). Bonnet and Wright (2014) recommended that the sample size for test-retest reliability should be at least 30 subjects to have a strong intercorrelation.

Subjects

For pre-testing and face validity process, 7 parents with normally developing children aged between 8 and 17 years old were recruited in the study. During the proofreading process, one lecturer from the Institut Pendidikan Guru (IPG) and one secondary school teacher who majored in Bahasa Melayu volunteered to be the proofreaders. During the pilot study, two groups consisted of parents of normally developing children (n = 15) and parents of children with a secondary diagnosis (n = 15). Both groups must have children aged between 8 and 17 years old, respectively. It can improve the generalizability of their findings by better understanding how the participants answer the Malay-5-15R questionnaire by comparing two groups. The findings between the groups are compared based on the answers given by the participants (Shadish & Campbell, 2002).

Materials

For this research study, the initial version of the Malay-5-15R by Zainal and Jusoh (2023), which has been content-validated, was used for further evaluations of its psychometric properties. The 181 items remain the same in the questionnaire.

Research Design

The quantitative and qualitative research designs were chosen for the research design. As for the pre-test and face validity, it has been measured quantitatively and qualitatively. These non-numerical observations are arranged, compiled, and interpreted using qualitative data analysis. The development of concepts that help understand the causes of occurrences in natural contexts, as opposed to experimental ones, is the aim of qualitative research. It provides an appropriate focus on the meanings, experiences, and opinions of each participant in the research study (Verhoef & Casebeer, 1997).

In the proofreading process, it has been measured qualitatively by gathering proofreaders' comments and suggestions for the content-validated questionnaire. As for the pilot study, it can be measured quantitatively using IBM Statistical Package for the Social Sciences (SPSS) Statistics 20 software for internal reliability and convergent-divergent validity testing.

Procedures

Pre-testing

Participants, who were seven parents of normal developmental children aged between 8 to 17 years old, were recruited through social media such as WhatsApp and Facebook. The initial version of Malay-5-15R was in the Google Form platform and distributed through social media. The aim of the pre-testing and face validity is to rate the quality of the questionnaire based on criteria such as readability, feasibility, wording clarity, layout, and style of the questionnaire (Venkitachalam, 2015). The participant will rate it as 'yes' (1 mark) and 'no' (0 mark) for each criterion.

Proofreading

After taking some considerations from the comments from the pre-testing and face validity, proofreading of the questionnaire is needed to identify any grammatical and wording errors used and the appropriateness of the sentence structure in the questionnaire. One-week period was given to the proofreaders to complete the proofreading process. After that, the researcher will make amendments from the suggestions from the proofreaders.

Pilot Study

A pilot study was conducted to examine its internal consistency reliability, test-retest reliability, and inter-item correlation (for convergent and divergent validity). A pilot study is needed as one of the validation processes to detect any issues or gaps in those questionnaires within a small group of the intended subjects (Gudmundsson, 2012). Additionally, it can evaluate the target population's comprehension and challenges with completing the questionnaire (Hassan et al., 2006).

Parents received questionnaires for the pilot study via online (Google Form) methods. The questionnaires were given to parents who resided close to the researcher's location, the International Islamic University Malaysia (IIUM) Hearing and Speech Clinic (HSC), Indera Mahkota Campus, for online distribution. The researcher had obtained permission from the clinic manager of the IIUM HSC to recruit subjects from the clinic via the online method. Online invitations and links to take part in the pilot study were distributed for the online questionnaire using two different channels: 1) social media (such as Facebook); and 2) WhatsApp and Telegram groups. Following their agreement to take part in the pilot study, parents were requested to complete the Malay-5-15R form after receiving the informed consent through Google Form. Instructions were given to the participants to complete the Malay-5-15R form. It is recommended by studies from Trochim and Donnelly (2008) that, for the second pilot study, the subjects had two weeks to complete the questionnaire once more. This two-week gap is considered a long enough timeframe to avoid recall effects (Trochim & Donnely, 2008). The questionnaire was sent back to the respondents via email and WhatsApp.

Data analysis

This study was conducted from pre-testing to the pilot study. All data were analysed descriptively (i.e., bar charts) and statistically (using Statistical Package for the Social Sciences version 20 (SPSS 20 Software). The proposed data analysis based on the objectives is summarised in Table 1:

According to George and Mallery (2010), Cronbach's α values for internal consistency reliability are considered acceptable when they are greater than 0.70. Values less than 0.70 are considered questionable and require further review. According to DeVon et al. (2007), for convergent validity, inter-item correlation values of more than 0.30 were considered sufficient and a significant correlation between the domains. The significant correlations show that the domains were measured to have a similar construct and support each other (Napper et al., 2008). For divergent validity, in contrast to convergent validity, Campbell and Fiske (1959) proposed that divergent validity could be demonstrated if the r-value for inter-item correlation was less than 0.85.

Table 1: Summary of Data Analysis According to the Specific Objectives

Objectives	Expected Data	Proposed Data Analysis
To determine the face validity of the Malay-5-15R	Scores of readabilities, feasibility, wording clarity, style for each item in the questionnaire; Comments and suggestions by the subjects	Descriptive data analysis
To examine the internal consistency reliability of the Malay-5-15R	The scores of Malay-5-15R of all subjects	Statistical data analysis - Cronbach's alpha method
To examine the test-retest reliability of the Malay-5-15R	The scores of Malay-5-15R of all subjects	Statistical data analysis - Intra-class correlation coefficient (ICC)
To examine the convergent and divergent validity within and between domains of the Malay-5-15R	The scores of Malay-5-15R of all subjects	Statistical data analysis - Inter-item correlation

RESULTS

Pre-testing and Face Validity

The initial version of the Malay-5-15R was pre-tested on the parents of children with normal developmental ages between 8 and 17 years old (n=7). The participants were asked to rate the items in this initial version based on their criteria of readability, feasibility, wording clarity, layout and style of the questionnaire using a yes or no method (Venkitachalam, 2015). The items with a percentage of 90% and above were considered acceptable in this study (Yusoff, 2019). For this study, the researcher found that all eight domains of the pre-testing score were acceptable, with a 100% score for readability, feasibility, and layout and style, while the score for clarity of words was 91.7%. The time taken to complete the initial version of the Malay-5-15R was longer, as reported by all participants, who took about 10 to 25 minutes to complete the questionnaire. One participant is given some suggestions for the questionnaire on the impact question parts (e.g., suggested on "Tidak Berkaitan" option for the impact question). Other than that, other participants reported that all items were understandable. Modification was made according to the comments and suggestions by participants, with the total number of items remaining 181.

Proofreading

Next, two Malay language experts proofread the initial version of the Malay-5-15R before it was pilot tested with the actual and intended population. Both experts have more than five years of experience teaching Bahasa Malaysia (BM); one teaches at SMK Tamparuli, Sabah, a secondary school, and the other is a lecturer at Institut Pendidikan Guru, Kampus Kent, Sabah. After a week of proofreading the questionnaire, a few minor grammatical mistakes were found in this initial version. The finalised version of the Malay-5-15R (Appendix A) was created as a result of the adjustments made. Table 2 shows some of the amendments from the initial version.

Table 2: Amendments from the initial version of Malay-5-15R

No	Items	Proofreader 1	Proofreader 2	Finalized version
27	Selalu dalam pergerakan berterusan atau sukar duduk diam (jari bergerak- gerak, mencabut sesuatu dan lain-lain)	Sentiasa bergerak atau sukar duduk diam (jari bergerak- gerak, mencabut sesuatu dan lain-lain)	Selalu bergerak secara berterusan atau sukar duduk diam (jari bergerak- gerak, mencabut sesuatu dan lain-lain)	Selalu bergerak secara berterusan atau sukar duduk diam (jari bergerak- gerak, mencabut sesuatu dan lain-lain)
61	Kesukaran mengingat maklumat tentang data peribadi, seperti tarikh lahir, alamat rumah dan sebagainya	Kesukaran untuk mengingat maklumat tentang data peribadi, seperti tarikh lahir, alamat rumah dan sebagainya	Kesukaran mengingat maklumat berkaitan data peribadi, seperti tarikh lahir, alamat rumah dan sebagainya	Kesukaran untuk mengingat maklumat berkaitan data peribadi, seperti tarikh lahir, alamat rumah dan sebagainya

Pilot Study

Thirty-five participants were recruited using the purposive sampling method for the pilot study. The inclusion criterion for this pilot study was a group of parents having typically developing children aged 8 to 17 years old (n=19). Another group was parents having children with secondary diagnoses aged 8 to 17 years old (n=16). All the responses and data were analysed for further statistical analysis using SPSS 20. In this pilot study, two different reliability analysis techniques were used. These were: i) the test-retest reliability was determined using the intra-class correlation coefficient (ICC); ii) the internal consistency of the items was assessed using the Cronbach's alpha method. From this pilot study, the convergent and divergent validity analysis was also calculated. Below is an explanation of the reliability and validity analyses' findings.

Internal Consistency Reliability Analysis

The coefficient alpha value showed the internal consistency of the initial version of the Malay-5-15R. Based on this result, it has been found that the initial version of Malay-5-15R has excellent values for each of its sub-domains ($\alpha = 0.93 – 0.97$) and an excellent overall α value of 0.99. Table 3 displays the specific values of α for every sub-domain. According to these findings, the initial version of the Malay-5-15R exhibited excellent internal consistency ($\alpha > 0.93$), indicating that it was prepared for validation testing against a broader sample.

Table 3: The Internal Consistency of the Initial Version of Malay-5-15R

Sub-domains	Coefficient alpha values (α)	Description
Motor skills	0.95	Excellent
Executive functions	0.97	Excellent
Perception	0.93	Excellent
Memory	0.95	Excellent
Language and communication	0.97	Excellent
Learning skills	0.96	Excellent
Social skills	0.97	Excellent
Mental health problems	0.97	Excellent

Test-retest Reliability Analysis

In the test-retest reliability process, the same subjects will answer the questionnaire again, with a two-week gap after the first pilot study**.** The consistency of the questionnaire's scoring over time was assessed using the intra-class correlation coefficient (ICC) analysis. The initial Malay-5-15R version demonstrated an excellent coefficient of stability in the ICC analysis, with a correlation coefficient of 0.98 (Williams, 2003). As for the coefficient stability, which was nearly 1, the result showed that each participant's score was closely associated with both pilot study 1 (P1) and pilot study 2 (P2). Each participant's responses in P2 were almost identical to those in P1. Thus, since the total item reliability of each domain and the scale was sufficient to proceed with other validation testing, no further modifications or amendments were required for the initial version of Malay-5-15R.

Convergent and Divergent Validity

The convergent and divergent validity were used for further validity testing. It is measured using inter-item correlation for validity testing. Most of the initial version of the Malay-5-15R items, as in Table 4, have a significant correlation with items in the sub-domains, which are 89% of items, with r-values of greater than 0.30 ($p < 0.05$), demonstrating convergent validity. The inter-item correlations of items from the Motor Skills domain were strongly correlated for divergent validity, as shown in Table 4, with r-values below the permitted value of 0.85 ($p < 0.05$) (Campbell and Fiske, 1959). As a result, the divergent validity of the Malay-5-15R for these domains was similarly acceptable. However, it is an exception for other sub-domains, whereby the r-values exceeded the recommended values, which were within 0.86-0.94. .

Table 4: The Summary of Convergent and Divergent Validity Through Inter-Item Correlation Analysis

Domains	Convergent Validity (r)	Divergent Validity (r)
Motor skills	0.05 - 0.80	-0.31 - 0.84
Executive functions	0.08 - 0.90	-0.20 - 0.86
Perception	-0.61 - 0.84	-0.21 - 0.87
Memory	0.31 - 0.85	-0.13 - 0.87
Language and communication	-0.37 - 0.94	-0.19 - 0.94
Learning skills	-0.60 - 0.92	-0.31 - 0.87
Social skills	0.04 - 0.91	-0.13 - 0.88
Mental health problems	-0.95 - 0.92	-0.16 - 0.94

DISCUSSION

The study was designed to describe the psychometric properties of the initial version of Malay-5-15R, which included reliability and validity testing. The initial version of Malay-5-15R needs further validation to be used among parents of the Malay-speaking population and other health professionals. For this study, the validation process was continued until the stages of convergent and divergent analysis.

Firstly, the pre-testing and face validity were conducted on small groups of parents with typically developing children aged 8 to 17 years old. Overall, all eight domains of the Malay-5-15R score were acceptable, with a 100% score for readability, feasibility, and layout and style, while the score for clarity of words was 91.71%. It is equivalent to the studies by Kadesjö et al. (2004), which reported the items from the 5-15R as highly relevant. However, the subjects reported that it takes almost 10 to 25 minutes to answer the 181 questions in Malay-5-15R. Both Galesic and Bosnjak (2009) and Sandelin (2022) report similar findings, as answering lengthy questionnaires has both advantages and disadvantages. Thus, providing subjects with additional time to answer a questionnaire allows them to think it through and provide a more accurate and comprehensive response. Length durations contribute to lower mistake rates and higher-quality data overall. Although this Malay-5-15R questionnaire is quite lengthy, according to Ziegler et al. (2014), long questionnaires can cover the entire domain of a construct, making them suitable for situations requiring a comprehensive assessment. Longer questionnaires are more likely to cover every aspect of a construct and can differentiate respondents across a wide variety of demographics, which results in a more comprehensive and deeper analysis. In contrast, longer questions appeared to have an adverse effect on response rates. Other than that, subjects may become uninterested as the long questionnaire carries on, which could result in fewer meaningful and engaging answers. However, as this questionnaire is still in the early stages of the validation process, all items were retained despite some parents expressing concerns about its length. This decision was made because further validation procedures, such as assessing construct validity, may be necessary to enhance its overall validity. Based on the findings from construct validity analysis, irrelevant items can be identified and subsequently removed to reduce the total number of items in the questionnaire.

After the pre-testing and face validity, the questionnaire was proofread before conducting a pilot study. Based on this study, only a few minor grammatical mistakes were found in this initial version. Other than that, there are no significant issues in this initial version, Malay-5-15R. Proofreading is needed for this study as the initial version of Malay-5-15R has been cross-culturally adapted and translated in the previous study. The role of proofreaders is to find any mistakes that were missed in the initial translation and fix them accordingly to suit the Malay-speaking population. This includes grammatical mistakes, incorrect terms, and sentence problems that potentially alter the meaning. This can enhance the quality of the questionnaire (Niki, 2015). Overall, the initial version of Malay-5-15R is suitable for use by the Malay-speaking population.

For internal reliability, the research conducted on children in the Swedish population by Kadesjö et al. (2004) and the Danish population by Lambek & Trillingsgaard (2014) reported acceptable to excellent Cronbach's alpha values for all subdomains. It is equivalent to the initial version of Malay-5-15R, which has excellent values for each of its sub-domains ($\alpha = 0.93 – 0.97$) and an excellent overall α value of 0.99. As the reliability is good, it produces similar scores of the Malay-5-15R when being tested repeatedly over time. Illum and Gradel (2014) obtained a similar result, which was tested with parents having neurological disabilities in children. It also reported an excellent Cronbach's value for all subdomains ($\alpha = 0.96$). Trillingsgaard et al. (2004) also conducted the clinical validity and utility of the FTF, and its resulting internal consistency was high, with a Cronbach's Alpha ranging from 0.84 to 0.93. However, for the Malaysian context, there is still a limited number of studies that relate to the usage of Malay-5-15R.

Next, other psychometric properties, which are the test-retest reliability of the Malay-5-15R, were also measured in this study by using ICC analysis. It is found that the Malay-5-15R have an excellent coefficient of stability using the ICC analysis, with a correlation coefficient value of 0.98. As a result, even after testing over a different period, it was concluded that this questionnaire was sufficiently stable and consistent. Kadesjö et al. (2004) also obtained a similar result, as the overall retest stability over time was good, which ranged from 0.74 to 0.91.

Furthermore, this study also aimed to confirm the convergent and divergent validity of the items in the Malay-5-15R by measuring the inter-item correlations. This study managed to obtain a strong correlation for most of the items (89%) in the sub-domains, with r-values of greater than 0.30, demonstrating that convergent validity was established. This indicated that most items within the

same sub-domains had strong correlations, suggesting that the Malay-5-15R sub-domain's convergent validity was acceptable. Good convergent validity means that Malay-5-15R measures the desired construct precisely and consistently, which is essential for a questionnaire. Besides, the items from each domain from Malay-5-15R are relevant and correlated with each other. Trillingsgaard et al. (2004) also conducted a validity study in which three of the four relevant FTF domains (e.g. language, learning and perception) were tested for convergent validity in relation to the Wechsler Intelligence Scale for Children - Third Edition (WISC-III). The study showed that language and learning domains in FTF have a significant correlation with WISC domains, with r-values between 0.38 and 0.47.

The divergent validity was also established for the Motor Skills domain, with r-values of less than 0.85, while the other domains exceeded the recommended values. About 0.01% of the items in Malay-5-15R were more than 0.85. This indicates that the items from those domains may not be distinct enough from each other. Other than that, high divergent validity might suggest that the constructs overlap with items from other domains, indicating they assess the same or related underlying concepts rather than distinct ones (Campbell & Fiske, 1959). Table 5 shows an example of items from the questionnaire that portray high divergent validity. From the example, the divergent validity value was 0.86. The item FE023 was strongly correlated with its own domain (Executive Function) rather than another domain (Memory). It is due to the r-values from its own domain being higher than those from the other domain. In contrast with ME070, it was strongly correlated to other domains (FE) rather than its own domain (ME) due to the lower r-values. Trillingsgaard et al. (2004) also conducted a validity testing for FTF questionnaire among Danish diagnostic groups. The study measures the convergent validity that refers to the ability of the FTF to distinguish between children with different diagnostic conditions and normal children. Based on the study, all the domains except for Mental Health Problems showed significant differences between the diagnostic groups and the norm sample, indicating good discriminant validity.

There are certain items that do not achieve acceptable divergent validity. Other psychometric properties may be beneficial as they can further examine the Malay-5-15R questionnaire. Besides, this can help to revise and review the items in the Malay-5-15R. Thus, it is recommended to examine its construct validity of the Malay-5-15R to determine whether the item is loading to the intended domains using factor analysis (DeVon, 2007).

Table 5: Example of Items Having High Divergent Validity

Items	Divergent (r-values)	Question	Items	Divergent (r-values)	Question
FE023	0.88	Sering mengelak, tidak suka, atau enggan melibatkan diri dalam tugasan yang memerlukan usaha mental (tumpuan pemikiran) yang berterusan (seperti kerja rumah)	ME070	0.79	Kesukaran mengingati arahan yang panjang atau arahan berbilang langkah
BA081	0.94	Kesukaran menjelaskan apa yang dia mahu	KS125	0.84	Kesukaran menerangkan emosi secara lisan apabila berasa kesunyian, bosan dan lain-lain

CONCLUSION

This study concluded that the Malay-5-15R questionnaire has good face validity, internal consistency reliability, test-retest reliability and convergent and divergent validity, which are important psychometric properties in a questionnaire. To ensure a good translation and adapted questionnaire, a validation and reliability process is required to make it useful to the Malay-speaking population. In Malaysia, there are still lacking assessment tools or rating scales that are translated and adapted into Malay, as there are specific questions that are not suitable and appropriate to our culture. Thus, the translated and validated version of Malay-5-15R will be one of the tools that can be used among the parents, teachers and professionals in the clinical setting in Malaysia, as it gives an impression of deficits in certain domains in Malay-5-15R.

LIMITATIONS AND FUTURE RECOMMENDATIONS

This study has a few limitations that should be addressed. Firstly, the process of this study is only completed until convergent and divergent validity. This is because the 181 items in Malay-5-15R affect how long it takes the subjects and the researcher to finish the pilot study part. Thus, other psychometric properties should be examined in future studies to produce the finalised version of Malay-5-15R. Further validation process may impact some of the items, which will produce a shorter and revised version of Malay-5-15R. This will shorten the parents' questionnaire completion time without limiting the amount of information required from them.

ACKNOWLEDGEMENT

The authors would like to express their deepest gratitude to everyone who helped this study be carried out successfully, either directly or indirectly. This research was not funded by any grant.

REFERENCES

Bohlin, G., & Janols, L. O. (2004). Behavioural problems and psychiatric symptoms in 5-13 year-old Swedish children - A comparison of parent ratings on the FTF (Five to Fifteen) with the ratings on CBCL (Child Behavior Checklist). *European Child and Adolescent Psychiatry, Supplement, 13*(3). https://doi.org/10.1007/s00787-004-3003-1

Bonett, D. G., and Wright, T. A (2014). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size. *Journal of Organizational Behaviour, 36*(1)

Camden C, Dostie R, Heguy L, et al. (2020). Understanding parental concerns related to their child's development and factors influencing their decisions to seek help from health care professionals: Results of a qualitative study. *Child Care Health Dev, 46*, 9–18. https://doi.org/ 10.1111/cch.12731

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105. doi:10.1037/h0046016

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Kostas-Polston, E. (2007). A Psychometric Toolbox for Testing Validity and Reliability. *Journal of Nursing Scholarship, 39*(2), 155-164. doi:10.1111/j.1547-5069.2007.00161.x

Galesic, M., & Bosnjak, M. (2009). Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly, 73*(2), 349–360. doi:10.1093/poq/nfp031

George, D., & Mallery, P. (2010). SPSS for Windows step by step: a simple guide and reference, 17.0 update (10th ed.). Boston: Allyn & Bacon.

González, C., Castillo, R. D., Franzani, J. P., & Martinich, C. (2021). Screening of developmental difficulties during the transition to primary school. *International Journal of Environmental Research and Public Health, 18*(8). https://doi.org/10.3390/ijerph18083958

Gudmundsson, E. (2012). Guidelines for translating and adapting psychological instruments. *Nordic Psychology, 61*(2), 29–45. https://doi.org/10.1027/1901-2276.61.2.29

Hassan, Z. A., Schattner, P., & Mazza, D. (2006). Doing A Pilot Study: Why Is It Essential? Malaysian Family Physician. *The Official Journal of the Academy of Family Physicians of Malaysia, 1*(2-3), 70–73. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC 4453116/

In, J. (2017). Introduction of a pilot study. *Korean Journal of Anesthesiology, 70*(6), 601–605. https://doi.org/10.4097/kjae.2017.70.6.601

Julious, S. A. (2005). Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics, 4*(4), 287-291. doi:10.1002/pst.185

Kadesjö, B., Janols, L. O., Korkman, M., Mickelsson, K., Strand, G., Trillingsgaard, A., & Gillberg, C. (2004). The FTF (Five to Fifteen): The development of a parent questionnaire for the assessment of ADHD and comorbid conditions. *European Child and Adolescent Psychiatry, Supplement, 13*(3). https://doi.org/10.1007/s00787-004-3002-2

Kadesjö, B., Janols. L-O, Korkman, M., Mickelsson, K., Strand, G., Trillingsgaard, A., Lambek, R., Øgrim, G., Bredesen, A. M., & Gillberg, C. (2017). Five-To-Fifteen-Revised (5-15R). Available at www.5-15.org

Khan, I., & Leventhal, B. L. (2020). Developmental delay. In: StatPearls. StatPearls Publishing, Treasure Island

(FL); PMID: 32965902.

Korkman, M., Jaakkola, M., Ahlroth, A., Pesonen, A. E., & Turunen, M. M. (2004). Screening of developmental disorders in five-year-olds using the FTF (Five to Fifteen) questionnaire: A validation study. *European Child and Adolescent Psychiatry, Supplement, 13*(3). https://doi.org/10.1007/s00787-004-3005-z

Lambek, R., & Trillingsgaard, A. (2015). Elaboration, validation and standardization of the five to fifteen (FTF) questionnaire in a Danish population sample. *Research in developmental disabilities, 38*, 161-170.

Lindblad, I., Gillberg, C., & Fernell, E. (2011). ADHD and other associated developmental problems in children with mild mental retardation. The use of the "Five-To-Fifteen" questionnaire in a population-based sample. *Research in Developmental Disabilities, 32*(6), 2805–2809. https://doi.org/10.1016/j.ridd.2011.05.026

Napper, L. E., Wood, M. M., Jaffe, A., Fisher, D. G., Reynolds, G. L., & Klahn, J. A. (2008). Convergent and discriminant validity of three measures of stage of change. *Psychol Addict Behav, 22*(3), 362-371. doi:10.1037/0893- 164X.22.3.362

Niki (2015). *Importance of Proofreading in Translation*. Niki's Int'l Ltd. Retrieved from https://nilservices.com/document-translation-proofreading/

Rai, N., & Thapa, B. (2015). A study on purposive sampling method in research. Kathmandu, *Kathmandu School of Law, 5*(1), 8-15.

Reynolds, N., Diamantopoulos, A., & Schlegelmilch, B. (1993). Pre-Testing in Questionnaire Design: A Review of the Literature and Suggestions for Further Research INTRODUCTION: WHAT IS PRETESTING. In International Journal of Market Research.

Sandelin, F. (2022). The effects of questionnaire length on response rate, non-response bias, and data quality. The SOM Institute, University of Gothenburg

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin.

Shaw, K. A., Bilder, D. A., McArthur, D., Ashley Robinson Williams, Amoakohene, E., Bakian, A. V., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Pas, E. T., Salinas, A., Warren, Z., Williams, S., Esler, A., Grzybowski, A., Ladd-Acosta, C. M., Patrick, M., Zahorodny, W., & Green, K. K. (2023). Early Identification of Autism Spectrum Disorder Among Children Aged 4 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. *Morbidity and Mortality Weekly Report. Surveillance Summaries, 72*(1), 1–15. https://doi.org/10.15585/mmwr.ss7201a1

Trillingsgaard, A., Damm, D., Sommer, S., Jepsen, J. R. M., Østergaard, O., Frydenberg, M., & Thomsen, P. H. (2004). Developmental profiles on the basis of the FTF (Five to Fifteen) questionnaire: Clinical validity and utility of the FTF in a child psychiatric sample. *European child & adolescent psychiatry, 13*, iii39-iii49.

Trochim, W. M. K., & Donnelly, J. P. (2008). *The Research Methods Knowledge Base* (3rd ed.). Atomic Dog Publishing.

Okoye, C., Obialo-Ibeawuchi, C. M., Obajeun, O. A., Sarwar, S., Tawfik, C., Madeeha Subhan Waleed, Asad Ullah Wasim, Mohamoud, I., Afolayan, A. Y., & Mbaezue, R. N. (2023). Early Diagnosis of Autism Spectrum Disorder: A Review and Analysis of the Risks and Benefits. Curēus. https://doi.org/10.7759/cureus.43226

Venkitachalam, R. (2015). Validity and Reliability of Questionnaires. Retrieved from https://www.slideshare.net/Venkitachalam/validity-and-reliability-ofquestionnaires

Verhoef, M. J., & Casebeer, A. (1997). Broadening Horizons: Integrating Quantitative and Qualitative Research. *The Canadian Journal of Infectious Diseases*, *8*(2), 65–66. https://doi.org/10.1155/1997/349145

Williams, Y. (2003). Test-retest Reliability Coefficient: Examples & Concept. Retrieved from https://study.com/academy/lesson/test-retest-reliabilitycoefficient-examples-lesson-quiz.html

Yusoff, M. S. B. (2019). ABC of Response Process Validation and Face Validity Index Calculation. *Education in Medicine Journal, 11*(3), 55–61. https://doi.org/10.21315/eimj2019.11.3.6

Zainal, N. A. & Jusoh, M. (2023). Adaptation of the Revised Five-To-Fifteen (5-15R) Questionnaire into Malay-

5-15R: preliminary Findings. *International Journal of Allied Health Sciences, 7*(5), 24-35.

Ziegler, M., Poropat, A., & Mell, J. (2014). Does the length of a questionnaire matter?. *Journal of Individual Differences*. doi: http://dx.doi.org/10.1027/1614-0001/a000147

Appendix A: Finalised Version of Malay-5-15R Questionnaire

5-15R Questionnaire: Evaluation of Development and Behaviours for Children and Adolescents

Arahan:
Tandakan pada kotak yang mengandungi kenyataan yang anda rasa paling sesuai dengan kelakuan anak anda dalam situasi harian berbanding dengan kanak-kanak sebaya mereka. Sila fokus kepada kelakuan anak anda dalam tempoh 6 bulan kebelakangan. Bagi mendapatkan gambaran yang tepat tentang kelakuan anak anda adalah penting untuk anda melengkapkan semua soal selidik ini.

Anda akan ditanya sekiranya kefungsian anak anda dalam pelbagai domain menimbulkan masalah dalam kehidupan seharian. Sila pertimbangkan sama ada masalah tersebut memberi kesan kepada anak anda dan orang lain di rumah, sekolah dan dalam kalangan rakan-rakan.

No.	**Item**	**Pemarkahan**		
		Tidak berlaku	**Kadangkala berlaku**	**Berlaku**
Kemahiran motor - kemahiran motor kasar; penggunaan badan anak dalam pelbagai aktiviti				
1.	Kesukaran menguasai kemahiran motor baharu seperti belajar menunggang basikal, bermain papan luncur, dan berenang			
2.	Kesukaran membaling dan menangkap bola			
3.	Kesukaran untuk berlari laju			
4.	Mengalami kesukaran atau tidak suka untuk menyertai sukan seperti bola sepak, hoki padang, dan bola keranjang			
5.	Masalah keseimbangan; contohnya kesukaran berdiri menggunakan satu kaki			
6.	Kerap tersadung dan jatuh			
7.	Cuai atau membuat pergerakan janggal			
Kemahiran motor - kemahiran motor halus; penggunaan tangan anak:				
8.	Tidak suka melukis, mengalami kesukaran melukis rajah yang menggambarkan sesuatu			
9.	Kesukaran mengendali, memasang dan memanipulasi objek kecil			
10.	Kesukaran menuang air ke dalam gelas tanpa tumpah			
11.	Kerap menumpahkan makanan ke atas pakaian atau meja ketika makan			
12.	Kesukaran menggunakan pisau, sudu dan garfu			
13.	Kesukaran mengenakan butang pada pakaian atau mengikat tali kasut			
14.	Kesukaran menggunakan pen (contoh: tekan terlalu kuat, tangan menggeletar)			

15.	Perkembangan tangan dominan yang tidak jelas, iaitu, tidak jelas sama ada dominan pada tangan kanan atau kidal			
16.	Sukar dan lambat untuk menulis			
17.	Pegangan pensel yang tidak kukuh, iaitu memegang pen dengan cara yang berbeza/luar biasa			
Jika anda memilih 'Kadangkala Berlaku' atau 'Berlaku' untuk soalan di atas, adakah masalah dengan fungsi motor mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Perhatian dan tumpuan: Keupayaan anak untuk memberi perhatian dan tumpuan terhadap pelbagai tugasan dan aktiviti:				
18.	Selalu gagal memberi perhatian sepenuhnya terhadap butiran atau membuat kesilapan kerana cuai			
19.	Selalu mengalami kesukaran untuk mengekalkan tumpuan terhadap tugasan atau aktiviti bermain			
20.	Selalu kelihatan seperti tidak mendengar apabila bercakap secara langsung dengannya			
21.	Mempunyai masalah untuk mengikut arahan dan gagal menyiapkan kerja sekolah, kerja rutin atau tugasan			
22.	Selalu mengalami kesukaran untuk mengatur tugasan dan aktiviti			
23.	Selalu mengelak, tidak suka, atau enggan melibatkan diri dalam tugasan yang memerlukan usaha mental (tumpuan pemikiran) yang berterusan (seperti kerja rumah)			
24.	Selalu kehilangan barang yang diperlukan untuk tugasan atau aktiviti (contoh, alat permainan, dan peralatan sekolah)			
25.	Selalu mudah terganggu dengan rangsangan luar (contoh, bunyi yang tidak relevan seperti orang lain bercakap, bunyi kereta yang dipandu)			
26.	Selalu terlupa sesuatu dalam melaksanakan aktiviti seharian			
Terlalu aktif dan impulsif; kecenderungan anak untuk menjadi terlalu aktif atau impulsif:				
27.	Selalu bergerak secara berterusan atau sukar duduk diam (jari bergerak- gerak, mencabut sesuatu dan lain-lain)			
28.	Kesukaran untuk kekal di tempat duduk (menggeliat di tempat duduk, kerap bangun dan bergerak)			
29.	Selalu berlari atau memanjat secara berlebihan dalam situasi yang tidak sesuai			

30.	Kesukaran untuk bermain dengan tenang dan senyap			
31.	Selalu “bergerak”atau sering bertindak seperti “tidak penat”			
32.	Selalu bercakap secara berlebihan/luar dari kebiasaan			
33.	Selalu menjawab soalan sebelum soalan selesai dikemukakan			
34.	Kesukaran untuk menunggu giliran (dalam permainan, semasa makan dan lain-lain)			
35.	Selalu mencelah atau mengganggu orang lain (contoh, mencelah perbualan atau permainan)			
Jika anda memilih ‘Kadangkala Berlaku’ atau ‘Berlaku’ untuk soalan di atas, adakah masalah dengan perhatian, tumpuan, terlalu aktif atau impulsif mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Kepasifan/tidak aktif: ketidakaktifan atau kecenderungan anak untuk menjadi terlalu pasif				
36.	Kesukaran untuk memulakan tugasan/aktiviti			
37.	Kesukaran untuk menyelesaikan tugasan/aktiviti, tidak menyiapkan tugasan/aktiviti seperti ahli kumpulan yang lain			
38.	Selalu “berada dalam dunia sendiri” atau mengelamun			
39.	Nampak pergerakan perlahan, tidak bermaya atau kurang tenaga			
Jika anda memilih ‘Kadangkala Berlaku’ atau ‘Berlaku’ untuk soalan di atas, adakah kepasifan atau tidak aktif mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Merancang/mengatur; keupayaan anak untuk merancang atau mengatur aktiviti				
40.	Kesukaran memahami akibat daripada tindakan sendiri (contoh, memanjat di tempat berbahaya, mematuhi aturan lalu lintas)			
41.	Kesukaran untuk merancang dan membuat persediaan untuk tugasan (contoh, mengumpul peralatan yang diperlukan untuk keluar atau ke sekolah)			
42.	Kesukaran untuk menyelesaikan tugasan berurutan (contoh, kanak-kanak kecil: memakai pakaian pada waktu pagi tanpa peringatan berterusan; kanak-kanak yang lebih besar: menyiapkan kerja rumah tanpa peringatan berterusan)			

Jika anda memilih 'Kadangkala Berlaku' atau 'Berlaku' untuk soalan di atas, adakah masalah untuk merancang/mengatur mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Persepsi terhadap ruang dan arah; persepsi anak terhadap ruang dan arah dalam dunia fizikal:				
43.	Kesukaran untuk mencari arah (walaupun di tempat yang mereka selalu berada)			
44.	Mudah terganggu dengan perbezaan ketinggian (walaupun sedikit) seperti ketika menaiki tangga dan sebagainya			
45.	Kesukaran menganggarkan jarak atau saiz			
46.	Kesukaran untuk memahami orientasi dan arah spatial (dimensi) (kanak-kanak kecil memusingkan pakaian bahagian belakang ke hadapan, kanak-kanak yang lebih besar keliru dengan huruf seperti b, p, d, atau nombor seperti 6, 9)			
47.	Mudah terlanggar orang lain apabila berada di tempat sempit			
Konsep masa; Keupayaan anak untuk memahami konsep masa:				
48.	Pemahaman tentang konsep masa yang lemah, contohnya, tidak mempunyai kemahiran untuk mengagak tentang berapa lama "lima minit" atau "satu jam" diambil atau tidak pasti tentang berapa lama sesuatu berlaku			
49.	Mempunyai idea yang samar-samar tentang waktu, sama ada pagi atau petang, sama ada masa untuk ke sekolah atau tidak			
50.	Berulang kali bertanya tentang tempoh masa sesuatu akan berlaku, contohnya, berapa lama lagi untuk keluar atau untuk ke sekolah			
51.	Boleh membaca jam mekanikal tetapi tidak memahami konsep masa yang sebenar			
Persepsi terhadap badan sendiri; persepsi anak terhadap tubuhnya sendiri dan tanggapan deria:				
52.	Tidak tahu sama ada pakaian muat atau tidak, tidak membetulkan stokin atau seluar yang melurut ke bawah			
53.	Persepsi yang lemah dan terlampau terhadap sejuk, sakit dan lain-lain			
54.	Kurang peka terhadap badan sendiri (tidak pasti saiz badan sendiri jika dibandingkan dengan persekitaran, contoh, terlanggar atau terjatuh ke atas sesuatu walaupun tanpa niat untuk berbuat demikian)			
55.	Terlalu sensitif terhadap sentuhan (merasa jengkel dengan pakaian yang ketat, menganggap sentuhan lembut sebagai kasar, dan lain-lain)			
56.	Kesukaran untuk meniru pergerakan orang lain			
Persepsi terhadap bentuk visual dan rajah; Keupayaan anak untuk menilai bentuk dan rajah:				

57.	Cenderung untuk salah tafsir gambar; contoh, mungkin menganggap gambar telur goreng sebagai gambar bunga			
58.	Kesukaran untuk mengenal pasti perbezaan yang tidak terlalu ketara dalam bentuk, rajah, perkataan dan corak yang seakan sama			
59.	Kesukaran untuk melukis gambar seperti gambar kereta, rumah dan lain-lain (berbanding dengan kanak-kanak sebaya)			
60.	Kesukaran untuk susun suai gambar			
Jika anda memilih 'Kadangkala Berlaku' atau 'Berlaku' untuk soalan di atas, adakah masalah dengan persepsi terhadap ruang dan arah, masa, badan sendiri, atau bentuk dan rajah mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Memori; Keupayaan anak untuk mengingat fakta atau apa yang dialami				
61.	Kesukaran untuk mengingat maklumat berkaitan data peribadi, seperti tarikh lahir, alamat rumah dan sebagainya			
62.	Kesukaran untuk mengingat nama orang lain (contoh, nama guru, rakan sekolah)			
63.	Kesukaran untuk mengingat nama hari dalam seminggu, bulan dan musim			
64.	Kesukaran untuk mengingat fakta bukan peribadi yang dipelajari di sekolah (contoh, peristiwa bersejarah, formula kimia dan lain-lain)			
65.	Kesukaran untuk mengingat perkara yang baru berlaku, seperti individu yang baru menelefon atau, apa-apa yang dimakan beberapa jam yang lalu dan lain-lain			
66.	Kesukaran untuk mengingat peristiwa yang berlaku suatu ketika dahulu, seperti peristiwa yang berlaku ketika melancong, hadiah hari jadi yang diperoleh dan lain-lain			
67.	Kesukaran untuk mengingat lokasi/tempat dia meletakkan sesuatu			
68.	Kesukaran untuk mengingati janji temu dengan rakan sebaya atau apa kerja rumah yang ada			
69.	Kesukaran untuk mempelajari rima, lagu, jadual pendaraban dan lain-lain dengan baik			
70.	Kesukaran untuk mengingati arahan yang panjang atau arahan pelbagai langkah			
71.	Kesukaran untuk menguasai kemahiran baharu, seperti peraturan permainan baharu			
Jika anda memilih 'Kadangkala Berlaku' atau 'Berlaku' untuk soalan di atas, adakah masalah dengan memori/ingatan mengganggu fungsi harian anak anda? Tidak berkaitan				

Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Pemahaman bahasa pertuturan; keupayaan anak untuk memahami bahasa dan pertuturan:				
72.	Kesukaran untuk memahami penerangan dan arahan			
73.	Kesukaran untuk mengikuti cerita yang dibaca dengan nada/suara yang kuat			
74.	Kesukaran untuk memahami apa-apa diperkatakan oleh orang lain (sering berkata “apa?” “apa maksud awak?”			
75.	Kesukaran untuk memahami konsep abstrak seperti “lusa”, “dalam turutan yang betul”			
76.	Cenderung untuk salah tafsir tentang sesuatu perkara yang diberitahu			
Bahasa ekspresif; keupayaan anak menghasilkan bahasa bahasa dan menyebut perkataan:				
77.	Bunyi pertuturan yang tidak pasti dan cenderung untuk salah menyebut perkataan			
78.	Kesukaran untuk mempelajari nama-nama warna, orang, huruf dan lain-lain			
79.	Kesukaran untuk mencari perkataan atau menjelaskan kepada orang lain, berkata: “itu, itu, itu… atau emm, emm, emm…”			
80.	Cenderung untuk mengingati perkataan dengan salah, menyebut “padang” dan bukannya “pandan”, merujuk kepada “tetikus” bukannya “penunjuk” dan lain-lain			
81.	Kesukaran menjelaskan apa-apa yang dia mahu			
82.	Kesukaran untuk bercakap dengan lancar tanpa tersekat-sekat			
83.	Kesukaran untuk mengekspresikan dirinya dalam bentuk ayat lengkap, dalam ayat yang betul dari segi tatabahasa, atau melakar maksud perkataan			
84.	Menyebut bunyi tertentu secara tidak betul (pelat menyebut ‘s’, kesukaran menyebut bunyi "r", suara sengau dan lain-lain)			
85.	Kesukaran untuk menyebut perkataan yang kompleks seperti “elektrik”, “infrastruktur” dan sebagainya			
86.	Mempunyai suara yang serak			
87.	Gagap			
88.	Bercakap dengan kadar yang sangat pantas sehingga sukar untuk orang lain memahami apa-apa yang dia katakan			
89.	Mempunyai pertuturan yang bercelaru/maksud yang tidak jelas			
Komunikasi lisan; keupayaan anak untuk menggunakan bahasa dan kemahiran untuk berkomunikasi dengan orang lain:				
90.	Kesukaran untuk menceritakan pengalaman atau situasi agar difahami oleh pendengar (contoh, peristiwa yang berlaku pada siang hari atau semasa cuti sekolah)			

91.	Kesukaran untuk mengekalkan topik perbualan apabila menceritakan sesuatu perkara kepada orang lain			
92.	Kesukaran untuk mengambil bahagian dalam perbualan, contoh, masalah beralih daripada mendengar kepada bercakap			
Jika anda memilih 'Kadangkala Berlaku' atau 'Berlaku' untuk soalan di atas, adakah masalah dengan pemahaman bahasa, penggunaan bahasa, atau komunikasi lisan mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Pemerolehan kemahiran akademik; jika anak anda berumur di bawah 8 tahun, sila ke item 122 Soalan yang berkaitan dengan pembelajaran anak boleh menjadi sukar bagi ibu bapa tanpa maklumat daripada guru. Namun begitu, sila cuba jawab soalan berikut berdasarkan apa yang anda tahu atau apa yang anda dengar daripada guru anak anda.				
Membaca, menulis, mengira (hanya kanak-kanak berumur 8 tahun ke atas):				
93.	Memperoleh kemahiran membaca adalah lebih sukar daripada yang dijangkakan jika mengambil kira keupayaannya mempelajari perkara lain			
94.	Mengalami kesukaran untuk memahami apa-apa yang dibacanya			
95.	Kesukaran untuk membaca dengan suara kuat pada kelajuan biasa (membaca terlalu perlahan, terlalu cepat, atau gagal membaca dengan lancar)			
96.	Tidak suka membaca (contoh, mengelak membaca buku)			
97.	Membuat tekaan semasa membaca			
98.	Kesukaran mengeja			
99.	Mengalami kesukaran untuk membentuk huruf dan sukar menulis dengan kemas			
100.	Kesukaran untuk mempersiap diri dalam aktiviti penulisan			
101.	Kesukaran untuk memperoleh kemahiran asas matematik (penambahan, penolakan; iaitu tambah, tolak)			
102.	Kesukaran untuk menyelesaikan masalah matematik yang diberikan dalam bentuk bertulis			
103.	Kesukaran untuk mempelajari dan mengaplikasikan pelbagai peraturan matematik			
104.	Kesukaran untuk mempelajari dan menggunakan jadual pendaraban			
105.	Kesukaran untuk mencongak			
Mempelajari perkara baharu dan mengaplikasikan pengetahuan di sekolah (hanya kanak-kanak berumur 8 tahun ke atas):				
106.	Kesukaran untuk memahami arahan lisan			
107.	Kesukaran untuk memahami atau menggunakan istilah abstrak, contoh, istilah yang berkaitan dengan saiz, isipadu, arah dimensi			

108.	Kesukaran untuk mengambil bahagian dalam perbincangan dengan kanak-kanak lain			
109.	Kesukaran untuk mempelajari fakta atau memperoleh pengetahuan tentang dunia sekeliling.			
110.	Pengetahuan atau kemahiran yang luar biasa dalam sesuatu bidang			
111.	Pandai dalam perkara artistik/seni atau praktikal (bermain alat, melukis, kerja pembinaan)			
Penyelesaian masalah di sekolah dan pendekatan kepada situasi pembelajaran baharu (hanya kanak-kanak berumur 8 tahun ke atas):				
112.	Kesukaran untuk merancang dan mengatur aktiviti, (contoh, susunan perkara yang perlu dilakukan, berapa banyak masa yang diperlukan untuk menguruskan tugasan tertentu)			
113.	Kesukaran untuk menukar rancangan atau strategi apabila diperlukan (contoh, apabila pendekatan awal gagal)			
114.	Kesukaran untuk memahami penjelasan dan mengikut arahan yang diberikan oleh orang dewasa			
115.	Kesukaran menyelesaikan tugasan abstrak (iaitu, bergantung kepada bahan pembelajaran yang boleh dilihat atau disentuh)			
116.	Kesukaran untuk terus mencuba dan menyelesaikan tugasan, sering meninggalkan tugasan separuh siap			
117.	Tidak bermotivasi untuk membuat kerja sekolah atau situasi pembelajaran yang setanding			
118.	Pembelajaran adalah lambat dan sukar			
119.	Melakukan sesuatu dengan cepat, terburu-buru atau tergesa-gesa			
120.	Tidak boleh/tidak akan bertanggungjawab atas tindakan sendiri, perlukan banyak pengawasan			
121.	Sangat memerlukan sokongan, yakni ingin tahu sama ada dia beraksi dengan baik			
Jika anda memilih 'Kadangkala Berlaku' atau 'Berlaku' untuk soalan di atas, adakah masalah akademik atau masalah pembelajaran mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Kemahiran sosial; keupayaan anak untuk mengambil bahagian dalam persekitaran sosial dan berinteraksi dengan orang lain:				
122.	Tidak memahami isyarat sosial orang lain, contoh, ekspresi muka, gerak isyarat, nada suara atau bahasa badan			
123.	Kesukaran untuk memahami perasaan orang lain			
124.	Kesukaran untuk bertindak balas terhadap keperluan orang lain (contoh, tidak menolong orang yang terjatuh)			

125.	Kesukaran untuk menerangkan emosi secara lisan apabila berasa kesunyian, bosan dan lain-lain			
126.	Bercakap dengan suara yang mendatar atau pelik			
127.	Kesukaran untuk menyatakan emosi dan reaksi dengan gerak isyarat muka atau bahasa badan			
128.	Bersikap/bergaya seperti orang yang jauh lebih berusia (tak sesuai dengan umur)			
129.	Kesukaran untuk berkelakuan seperti yang diharapkan oleh rakan sebaya			
130.	Kesukaran untuk menyedari cara berkelakuan dalam situasi sosial yang berbeza, seperti ketika melawat saudara-mara bersama ibu bapa, ketika melawat rakan-rakan, berjumpa doktor, pergi ke pawagam, dan lain-lain			
131.	Dianggap berbeza, pelik, atau tidak normal oleh rakan sebaya			
132.	Secara tidak sengaja memalukan diri sendiri sehingga ibu bapa berasa malu atau rakan sebaya mula tertawa			
133.	Sering seperti kurang akal budi *(common sense)*			
134.	Kesukaran untuk memahami jenaka			
135.	Memberi komen tanpa berfikir yang tidak sesuai dengan keadaan semasa			
136.	Kesukaran untuk memahami peraturan atau larangan			
137.	Sering bergaduh dengan rakan sebaya			
138.	Kesukaran untuk memahami dan menghormati hak orang lain, sebagai contoh, situasi di mana kanak-kanak kecil perlu bantuan lebih berbanding yang lebih besar, dan memahami permintaan ibu bapa untuk bersendirian, mengambil atau menggunakan barang orang lain tanpa izin dan lain-lain			
139.	Kesukaran dalam aktiviti atau permainan berkumpulan atau berpasukan, mencipta peraturan baharu untuk kepentingan sendiri			
140.	Kesukaran berkawan			
141.	Jarang berinteraksi dengan rakan sebaya			
142.	Kesukaran untuk menyertai aktiviti berkumpulan			
143.	Tidak diterima oleh kanak-kanak lain untuk menyertai permainan mereka			
144.	Tidak memperdulikan sentuhan fizikal seperti pelukan			
145.	Mempunyai satu atau beberapa minat yang memerlukan masa yang agak lama untuk terbentuk dan menjejaskan hubungan dengan keluarga dan rakan-rakan			
146.	Mengulang atau terperangkap dalam tingkah laku atau aktiviti yang kelihatan tidak bermakna			
147.	Berasa sangat kecewa dengan perubahan kecil dalam rutin harian			
148.	Hubungan mata (*eye contact*) yang tidak normal atau tiada semasa bersemuka			

Jika anda memilih 'Kadangkala Berlaku' atau 'Berlaku' untuk soalan di atas, adakah masalah dengan kemahiran sosial mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Masalah emosi:				
149.	Kurang keyakinan diri			
150.	Nampak seolah-olah tidak gembira, sedih, tertekan			
151.	Sering mengadu mengenai perasaan sunyi			
152.	Cuba untuk menyebabkan kecederaan diri sendiri atau bercakap tentangnya			
153.	Mempunyai kurang selera makan			
154.	Sering meluahkan perasaan tentang diri tidak berharga atau rendah diri berbanding kanak-kanak lain			
155.	Sering mengadu tentang sakit perut, sakit kepala, kesukaran bernafas atau gejala badan yang lain			
156.	Kelihatan tegang dan cemas atau mengadu tentang gugup/gementar			
157.	Menjadi sangat cemas atau tidak gembira apabila meninggalkan rumah, contoh, semasa pergi ke sekolah			
158.	Lebih banyak masalah tidur berbanding kebanyakan kanak-kanak yang sebaya			
159.	Sering mengalami mimpi ngeri			
160.	Berjalan dalam tidur atau mengalami serangan gangguan tidur pada waktu malam apabila dia tidak dapat didekati atau ditenangkan			
161.	Sering hilang sabar			
162.	Sering bergaduh/berdebat dengan orang dewasa			
163.	Sering enggan mengikut arahan orang dewasa			
164.	Sering mengusik orang lain dengan melakukan perkara yang dianggap sebagai provokasi			
165.	Sering menyalahkan orang lain atas kesilapan atau tindakan buruk diri sendiri			
166.	Mudah tersinggung, atau terganggu dengan orang lain			
167.	Sering terlibat dalam pergaduhan			
168.	Kejam terhadap haiwan			
169.	Berbohong dan menipu			
170.	Mencuri barang di rumah			

171.	Sering memusnahkan barang kepunyaan ahli keluarga yang lain atau kanak-kanak lain			
172.	Mengalami episod tahap aktiviti yang sangat tinggi dan idea yang banyak, yang mana episod ini berulang selama beberapa hari			
173.	Berulangkali berasa tidak sihat, lemah badan, tidak bertenaga atau mudah marah yang jelas kelihatan			
Tindakan atau pemikiran obsesif; Tindakan atau pemikiran yang tidak dapat dikawal				
174.	Mengulangi beberapa aktiviti secara berterusan atau mempunyai tabiat yang sangat sukar untuk diubah			
175.	Mempunyai idea yang obsesif/tetap			
176.	Mempunyai pergerakan yang tidak dapat dikawal, pergerakan tiba-tiba, kekejangan atau kerutan muka			
177.	Mengulangi pergerakan yang tidak bermakna, seperti menggelengkan kepala, menyentak badan dan menggendang jari			
178.	Mengeluarkan bunyi yang tidak bermakna seperti berdehem, bersin, menelan, menyalak, menjerit dan lain-lain			
179.	Kesukaran mendiamkan diri, contoh, bersiul, berdengung, menggumam			
180.	Mengulang perkataan atau sebahagian perkataan tanpa maksud			
181.	Menggunakan perkataan atau bahasa kesat dengan cara yang berlebihan/melampau			
Jika anda memilih ‘Kadangkala Berlaku’ atau ‘Berlaku’ untuk soalan di atas, adakah masalah emosi, tindakan atau pemikiran obsesif mengganggu fungsi harian anak anda? Tidak berkaitan Tidak sama sekali Sedikit Agak banyak Sangat banyak				
Terangkan masalah anak anda yang paling anda risaukan:				
Terangkan kelebihan anak anda:				

ORIGINAL ARTICLE

An Exploratory Corpus Study of Malay Word Structures in School-Aged Children

Fatin Syatirah Binti Fakrul Razi [1], Saiful Adli Bin Jamaluddin [2], Wan Aslynn Wan Ahmad [2,*]

[1]ReGen Rehab Hospital, 46200 Petaling Jaya, Selangor, Malaysia
[2]Department of Audiology & Speech-Language Pathology Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

ABSTRACT

Introduction: Corpus-based research plays an essential role in understanding language structures relevant to education and clinical practice. This study provides an initial exploration of Malay word structures in school-aged children, focusing on grapheme-phoneme correspondences, syllable patterns, and inflectional morphemes. The findings aim to support linguistic research and inform speech-language pathology interventions. **Methods:** A descriptive analysis was conducted on a paediatric Malay corpus comprising the 1,000 most frequent words used by school-aged children. The words were analysed in terms of grapheme-phoneme relationships, syllable structures, and inflectional morphemes. Data were summarised descriptively. **Results:** The majority of words demonstrated direct grapheme-phoneme correspondence. Disyllabic words (50%) and CV syllable structures were the most frequent patterns identified (47%). Inflectional analysis revealed that root words (65%) predominated, with relatively fewer affixed or reduplicated forms. **Conclusion:** This preliminary study highlights key features of Malay word structures in school-aged children, confirming the language's transparent orthography and prevalence of simple syllabic forms. These findings provide a foundation for future research and have potential applications in educational resources and speech-language pathology, particularly in developing assessment and intervention tools.

Keywords:
Malay corpora; word structures; school-aged children; grapheme-phoneme correspondence; morphology

INTRODUCTION

A corpus is a collection of texts, conversations, or speeches used to describe and analyse a language (Sinclair, 2005). In Malaysia, the Institute of Language and Literature (Dewan Bahasa dan Pustaka) provides an online database (http://sbmb.dbp.gov.my/) that enables the construction of Malay corpora from a wide range of sources, including newspapers, novels, and magazines. This database offers flexibility, as corpora can be generated based on different text types, publication sources, or time periods, thus serving as a valuable tool for both linguistic research and educational practice.

The Malay corpus helps in constructing tests and assessments. For example, Jamaluddin (2016), used the most frequent Malay words in a text corpus in developing the Malay Matrix Sentence Test, one of the speech tests in audiological evaluation. Additionally, speech corpus helps in the study of Malay speakers with speech impairments (Rosdi et al., 2017). Rosdi et al. (2017) conducted a study that consisted of a speech corpus for 30 speech impaired children with different diagnosis and severity level and measure the intelligibility of speech production.

For Malay language, Zakaria et al. (2021), generated 1000 most frequent words used in Paediatric Malay Corpora specifically for school-aged children. All words have different word structures such as the grapheme-phoneme correspondence.

The corpora for adults and children typically differ, as the most common words used in daily communication are not the same across age groups in Malaysia. Zakaria et al.

* *Corresponding author.*
E-mail address: wanaslynn@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

(2021) for instance, demonstrated that the phoneme distribution in the first 1,000 most frequent Malay words used by children differs substantially from that reported by for adults in Jamaluddin (2016). These differences are evident in vocabulary selection, with adult corpora containing more complex and varied word forms, while children's corpora reflect simpler lexical patterns. Farwell (2009) further notes that children's language is generally shorter and less complex than that of adults, which tends to involve greater variation and ambiguity.

The Malay Word Structures

Malay exhibits a transparent orthography and relatively complex syllable structures compared to languages such as English (Lee et al., 2013). The language contains four basic syllable types: V, VC, CV, and CVC, where 'V' represents a vowel and 'C' represents a consonant. The word *emak* /ə.mak/ (mother) consists of two syllables. The first syllable follows a V structure, while the second syllable follows a CVC structure. For the word *amboi* /am.boi/ (an utterance of disbelief or surprise), there are two syllables, which are VC and CVC respectively. Loanwords, however, often introduce more complex forms such as CCV structures (Sariyan, 2004, as cited in Lee et al., 2014). An example of a loanword in Malay *anggur* /aŋgur/ (grapes, borrowed from Persian) in which there is a consonant cluster CCV in the word.

Syllabic and phonetic structures are central to characterising Malay words (Lee et al., 2011). For instance, words of varying lengths may display distinct syllabic patterns, with long-string and short-string words showing different structural configurations. Moreover, the language's agglutinative nature means that affixation plays a key role in extending meaning. Affixes in Malay occur in several forms; prefixes, suffixes, circumfixes, and infixes; which systematically modify root words to create new grammatical or semantic functions (Benjamin, 2009).

Text corpora have been widely recognised across disciplines as valuable sources of linguistic and educational data (Awang, 2020). For example, Lee et al. (2013) analysed Malay word structures in children's storybooks and compared them with English. However, there remains limited research focusing specifically on the words most frequently used by school-aged children. To address this gap, the present study examines word structures in a paediatric Malay corpus.

The primary aim of this study is to analyse the structure of Malay words in school-aged children's for 7-12 years of age, corpora compiled by Zakaria et al. (2021). This work is particularly relevant to the field of speech-language pathology (SLP) in Malaysia, as it provides insights that can help identify, elicit, and correct morphological errors in children with speech and language difficulties. Since children with language impairments often produce incorrect word structures, understanding the most frequent forms used by their peers can guide clinicians in recognising and addressing such errors. Previous findings have highlighted the value of corpora in clinical applications, for instance, Mazenan et al. (2014) demonstrated the use of Malay corpora in articulation disorder interventions, while Carlisle et al. (2010, as cited in Lee et al., 2014) noted the role of morphological knowledge in reading comprehension.

Accordingly, this study pursues four specific objectives:

1. To identify grapheme-phoneme correspondences.
2. To examine syllable counts and structures within the corpus.
3. To describe the types of inflectional morphemes; prefixes, suffixes, circumfixes, reduplications, and irregular forms; present in the corpus.

MATERIALS AND METHODS

This study employed a descriptive design to analyse the structure of Malay words in a paediatric corpus. The corpus, developed by Zakaria et al. (2021), comprises the 1,000 most frequent words used by school-aged children. These words were examined for grapheme-phoneme correspondences, syllabic structures, and inflectional morphemes.

The Corpus

The corpus (Zakaria et. Al, 2021) was derived from both written and spoken sources of the Malay language. The written component was extracted from Malaysian primary school textbooks, and the 1,000 most frequent words were identified using the frequency of occurrence function. The spoken component was compiled from popular Malay animated films and television series, including *Geng: Pengembaraan Bermula*, *Boboiboy: The Movie*, *Upin dan Ipin*, and *Ejen Ali*, after which all dialogue was transcribed into written text.

To construct the final Paediatric Malay Corpus, the 1,000 most frequent words from the written and spoken sources were tabulated, cross-checked, and merged into a single list. Redundant entries were removed, and the frequency of each word was retained to ensure representativeness.

Criteria for Study Characteristics

Only Malay words, fillers, and exclamations were retained in the corpus, as these categories reflect the natural language use of school-aged children. In contrast, English words and abbreviations or short forms were excluded to maintain the linguistic focus on Malay and to avoid distortions arising from code-switching or non-standard orthography.

Research Procedure

The analysis of word structures in the paediatric Malay corpus was carried out at three linguistic levels: grapheme-phoneme correspondence, syllabic structure, and inflectional morphemes.

Grapheme-Phoneme Correspondence

Grapheme-phoneme correspondence was examined to determine the relationship between written symbols and spoken sounds. A phoneme is defined as the smallest linguistically distinctive unit of sound, whereas a grapheme represents its written form (Tan et al., 2012). For each word in the corpus, the number of letters and phonemes was calculated. Words were then classified into two categories: (1) those with a direct one-to-one correspondence between graphemes and phonemes, and (2) those without such direct correspondence.

Syllabic Structure

The number of syllables in each word was calculated to identify overall syllabic patterns within the paediatric Malay corpus. For example, the word *kita* (we) consists of two syllables, whereas *aktiviti* (activity) comprises four. The structural composition of syllables was also analysed, such as the V.CVC pattern found in the word *ikan*. This analysis enabled the identification of the most frequent syllable types and distributions.

Inflectional Morphemes

The analysis of inflectional morphemes was conducted to determine the grammatical modifications of root words. Inflectional morphemes, such as prefixes, suffixes, circumfixes, and reduplications, were identified and categorised. The relative frequencies of root words, prefix–root word combinations, root word–suffix combinations, circumfixed forms, and reduplicated words were calculated. This analysis provided insight into the distribution and prevalence of morphological structures within the corpus. Figure 1 shows the flowchart of the procedure carried out.

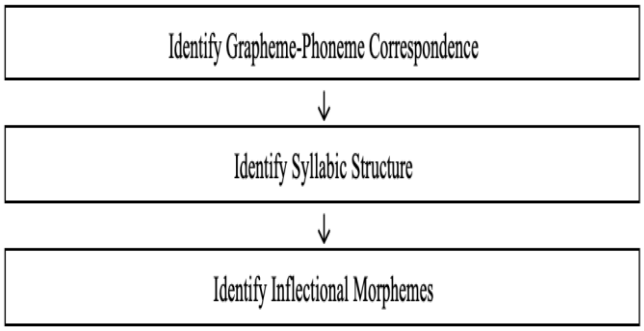


Figure 1: The Research Flowchart

RESULTS

The study data were analysed and presented descriptively. In addition, the results were organised into a summary table to provide a clear overview of the findings in line with the stated objectives and research questions.

Grapheme-Phoneme Correspondence

A total of 1,000 words from the paediatric Malay corpus (Zakaria et al., 2021) were analysed. As shown in Figure 2, 86% of the words displayed a direct one-to-one correspondence between graphemes and phonemes, reflecting the high transparency of Malay orthography.

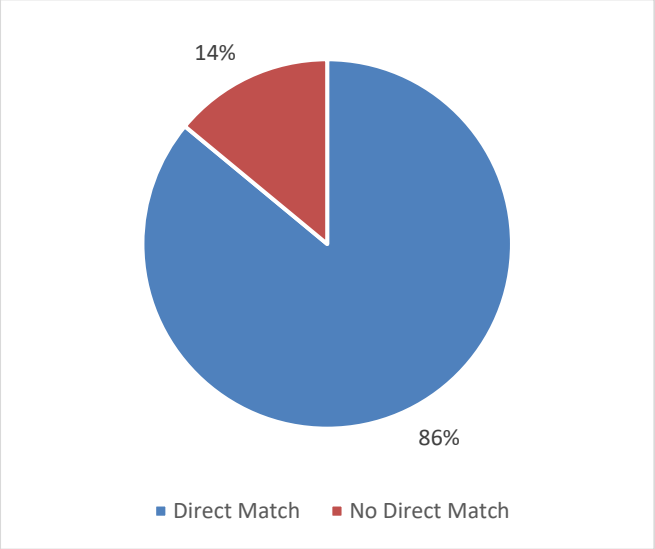


Figure 2. Grapheme-Phoneme Correspondence in the Corpus

Syllabic Structures

Syllabic structures were examined in two aspects: the number of syllables and the types of syllable patterns. As indicated in Figure 3, disyllabic words were the most frequent (50%), whereas words with five syllables represented the smallest proportion (3%).

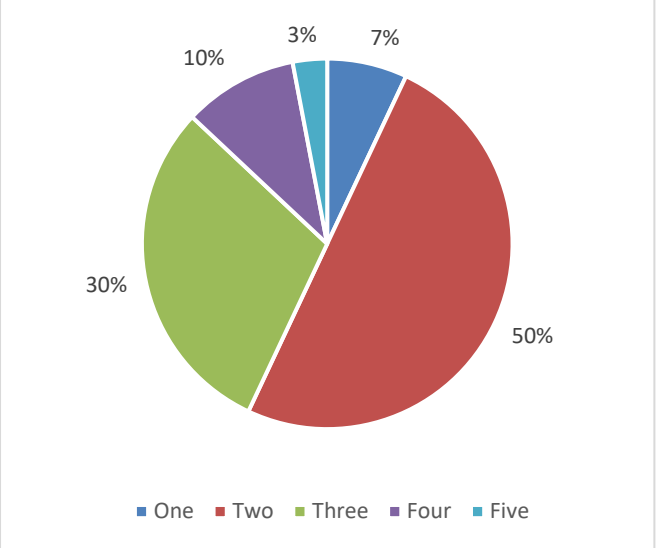


Figure 3. Distribution of Number of Syllables in the Corpus

Analysis of syllable types (Figure 4) revealed that the CV structure was the most common (47%), while the VC structure was the least common (7%).

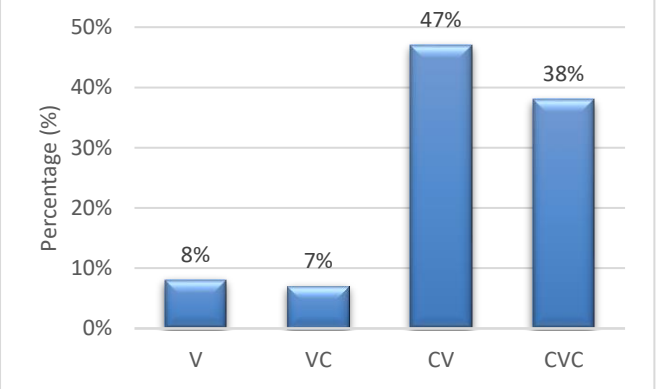


Figure 4. Distribution of Syllable Types in the Corpus

Inflectional Morphemes

Inflectional morphemes in the corpus included prefixes (e.g., *ber+jaya*), suffixes (e.g., *makan+an*), circumfixes (e.g., *ber+dasar+kan*), reduplications (e.g., *gambar-gambar*), and affixed reduplications (e.g., *men+jerit-jerit*). As presented in Figure 5, root words accounted for the majority of cases (65%), while reduplicated words were the least frequent (3%).

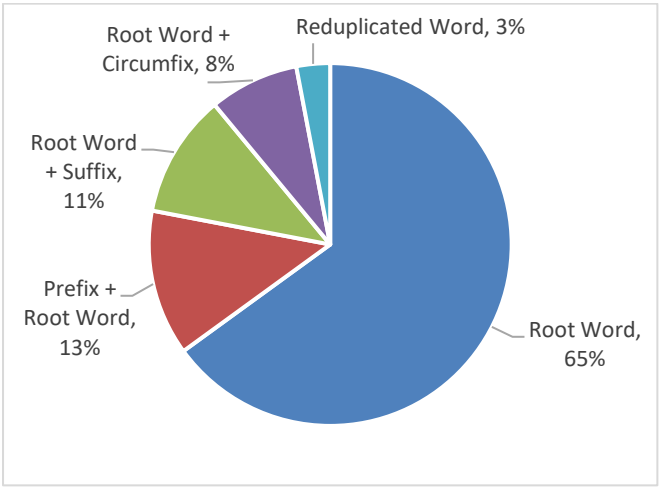


Figure 5. Distribution of Inflectional Morphemes in the Corpus

Overall Findings

Across all three analyses, the corpus revealed strong grapheme-phoneme transparency, a predominance of disyllabic words with CV structures, and a higher proportion of root words compared to affixed or reduplicated forms. These findings highlight the regularity and simplicity of Malay word structures commonly used by school-aged children.

DISCUSSION

This study conducted a descriptive analysis of word structures in the paediatric Malay corpus, focusing on grapheme-phoneme correspondences, syllabic structures, and inflectional morphemes. The results revealed clear patterns that align with initial assumptions: a predominance of words with direct grapheme-phoneme correspondences, a higher proportion of disyllabic words with CV syllable structures, and a greater frequency of root words compared to affixed forms.

Grapheme-Phoneme Correspondence

The analysis of grapheme-phoneme correspondences highlights the transparency of Malay orthography, where letters typically map directly onto sounds. This finding is consistent with Lee et al. (2013), who also reported strong grapheme-phoneme alignment in Malay. Such transparency supports word recognition and fluency, making Malay relatively accessible for early reading

acquisition. Previous research has shown that grapheme-phoneme knowledge is a strong predictor of decoding ability and reading fluency (Nation & Hulme, 1997; Vernon, 1993). Beal-Alvarez (2011) further demonstrated that grapheme-phoneme awareness is a critical early reading skill, predicting variance in children's literacy outcomes (as cited in Anthony & Lonigan, 2004; Lonigan, Burgess, & Anthony, 2000). For clinical practice, these findings provide important implications for Speech-Language Pathologists (SLPs). For example, children with speech sound disorders may produce errors such as final consonant deletion (*makan* → /maka/), which can be identified and addressed by comparing with the expected grapheme-phoneme structure.xxx

Syllabic Structures

The syllabic analysis revealed that most Malay words in the corpus are disyllabic and structured around CV patterns, reflecting the typical phonological organisation of the language. This supports earlier findings by Lee et al. (2013). Syllable awareness is a key component of phonological awareness, which plays an important role in literacy development (Ott, 1997; Wright & Jacobs, 2003). For SLPs, accurate knowledge of syllable boundaries is essential for detecting errors in children's speech production. Moreover, the prevalence of CV structures suggests that syllabic features should be integrated into early reading instruction. As Lee et al. (2013) noted, incorporating disyllabic word stimuli in intervention programmes enhances children's decoding skills and supports phonemic manipulation. These findings underscore the importance of addressing both syllabic and phonemic awareness in therapy and instructional contexts.

Inflectional Morphemes

The analysis of inflectional morphemes showed that root words predominated in the corpus, with fewer instances of affixed or reduplicated forms. This indicates limited morphological variation in the vocabulary most frequently used by school-aged children. Morphological knowledge is vital for spelling and literacy development. Pacton et al. (2013) emphasised the principle of root consistency, whereby roots often retain their spelling across related forms. For example, *makan* serves as the base for *makanan*. Strong morphological instruction from the early grades can therefore reinforce spelling accuracy and vocabulary expansion (Apel & Laurence, 2011). For SLPs, understanding morphological structures is crucial in assessing and intervening with children who present with morphological or language-related difficulties. As children are expected to manipulate word structures across different contexts, morphological awareness becomes an important tool for both reading and language learning. Deacon et al. (2017) provided further evidence that children's reading development is sensitive to the morphological structure of derivative words.

Overall Implications

Taken together, the findings demonstrate that Malay word structures at the graphemic, syllabic, and morphological levels contribute to literacy development and provide a foundation for clinical application in speech-language pathology. By highlighting the transparency of Malay orthography, the prominence of CV syllables, and the importance of root word morphology, this study underscores the need to integrate word structure knowledge into both educational and therapeutic practices for school-aged children.

CONCLUSION

This study analysed Malay word structures in school-aged children's corpora through three dimensions: grapheme-phoneme correspondences, syllabic structures, and inflectional morphemes. The findings confirmed the transparency of Malay orthography, the predominance of disyllabic CV structures, and the high frequency of root words relative to affixed forms. These results highlight the key orthographic and phonological features of Malay and provide practical implications for (SLPs. Specifically, knowledge of common word structures can assist SLPs in distinguishing between typical and erroneous productions, thereby improving intervention for children with speech and language difficulties. Early identification and remediation are crucial, as challenges in word structure can affect children's language, vocabulary, speech development, and reading ability.

Limitations and Suggestions

This study has several limitations. First, the analysis was based solely on the school-aged Malay corpora generated by Zakaria et al. (2021), without incorporating additional sources such as spontaneous speech samples directly produced by children. As a result, the corpus may not fully capture the natural variation present in children's everyday language use. Second, the analysis was time-intensive and required careful manual verification, which may have constrained the depth of exploration.

For future research, it is recommended that an updated version of the school-aged Malay corpora be developed using a broader range of sources, including children's

spoken data, as well as written materials such as newspapers, books, and magazines. A more diverse and representative corpus would provide a richer basis for understanding Malay word structures and better reflect children's actual linguistic output. Such advancements would strengthen both linguistic research and its clinical applications in speech-language pathology.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Abu Bakar, J., Khairuddin, K., Faidzul Nasrudin, M., & Zamri Murah, M. (2018). Pos-tagging malay corpus: A novel approach based on maximum entropy. *International Journal of Engineering & Technology*, *7(3.20),* 6. https://doi.org/10.14419/ijet.v7i3.20.18721

Alias, S., Mohammad, S. K., Keng Hoon, G., & Tien Ping, T. (2016). A Malay text corpus analysis for sentence compression using pattern-growth method. *Jurnal Teknologi, 78(8).* https://doi.org/10.11113/jt.v78.7413

Anderson, C. (2018). 6.3 inflectional morphology. Essentials of Linguistics. https://pressbooks.pub/essentialsoflinguistics/chapter/6-4-inflectional-morphology/#:~:text=Inflectional%20morphemes%20are%20morphemes%20that,a%20noun%20is%20inflectional%20morphology.

Awal, N. M., Ho-Abdullah, I., & Zainudin, I. S. (2014). Parallel corpus as a tool in teaching translation: Translating English phrasal verbs into Malay. *Procedia - Social and Behavioral Sciences, 112,* 882–887. https://doi.org/10.1016/j.sbspro.2014.01.1245

Bakar, N. S. (2020). The development of an integrated corpus for Malay language. *Lecture Notes in Electrical Engineering,* 425–433. https://doi.org/10.1007/978-981-15-0058-9_41

Beal-Alvarez, J. S., Lederberg, A. R., & Easterbrooks, S. R. (2011). Grapheme-phoneme acquisition of deaf preschoolers. Journal of Deaf Studies and Deaf Education, 17(1), 39–60. https://doi.org/10.1093/deafed/enr030

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112(4),* 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Chung, S.-F. (2011). Uses of ter- in Malay: A corpus-based study. *Journal of Pragmatics, 43(3),* 799–813. https://doi.org/10.1016/j.pragma.2010.10.004

Cohen-Mimran, R., Reznik-Nevet, L., Gott, D., & Share, D. L. (2022). Preschool morphological awareness contributes to word reading at the very earliest stages of learning to read in a transparent orthography. Reading and Writing. https://doi.org/10.1007/s11145-022-10340-z

Deacon, S. H., & Francis, K. A. (2017). How children become sensitive to the morphological structure of the words that they read. Frontiers in Psychology, 8. https://doi.org/10.3389/fpsyg.2017.01469

Farwell, C. B. (2009). The language spoken to children. Karger Publishers. https://karger.com/hde/article-abstract/18/4/288/156207/The-Language-Spoken-to-Children?redirectedFrom=fulltext

Harris, A. D., McGregor, J. C., Perencevich, E. N., Furuno, J. P., Zhu, J., Peterson, D. E., & Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in Medical Informatics. *Journal of the American Medical Informatics Association, 13(1),* 16–23. https://doi.org/10.1197/jamia.m1749

Ibrahim, B., Yunus, K., & Ibrahim, B. (2018). Perspectives on corpus linguistics: The methodological synergy in Second language pedagogy and research. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3258776

Jamaluddin, S. A. (2016). Development and evaluation of the digit triplet and auditory visual matrix sentence tests in Malay.

Lee, L. W., Low, H. M., & Mohamed, A. R. (2012). Word count analysis of Malay language textbooks for the purpose of developing a malay reading remedial programme. *Writing Systems Research, 4(1),* 103–119. https://doi.org/10.1080/17586801.2012.690713

Lee, L. W., Low, H. M., & Mohamed, A. (2013). A

Comparative Analysis of Word Structures in Malay and English Children's Stories. Undefined. https://www.semanticscholar.org/paper/A-Comparative-Analysis-of-Word-Structures-in-Malay-Lee-Low/902612e86018adf08602df263ec949daee5f42f6

Lee, L. W., & Low, H. M. (2014). Analysis of Malay word structure by pre-service special education teachers: Foundation-Level Knowledge for remedial instruction. *Australian Journal of Learning Difficulties, 19(1),* 33–46. https://doi.org/10.1080/19404158.2014.891531

Leech, G. N. (2011). Principles and applications of corpus linguistics. *Perspectives on Corpus Linguistics,* 155–170. https://doi.org/10.1075/scl.48.10lee

Louleli, N., Hämäläinen, J. A., Nieminen, L., Parviainen, T., & Leppänen, P. H. T. (2020). Dynamics of morphological processing in pre-school children with and without familial risk for dyslexia. Journal of Neurolinguistics, 56, 100931. https://doi.org/10.1016/j.jneuroling.2020.100931

Maseri, M., & Mamat, M. (2018). Malay language speech recognition for preschool children using Hidden Markov model (HMM) system training. *Lecture Notes in Electrical Engineering,* 205–214. https://doi.org/10.1007/978-981-13-2622-6_21

Mazenan, M. N., Swee, T. T., & Soh, S. S. (2014). Recognition test on highly newly robust Malay corpus based on statistical analysis for Malay articulation disorder. *The 7th 2014 Biomedical Engineering International Conference.* https://doi.org/10.1109/bmeicon.2014.7017394

Omar, S., Bakar, J. A., Mohd Nadzir, M., Harun, N. H., & Yusoff, N. (2021). Text simplification for malay corpus: A Review. *2021 International Conference on Computer & Information Sciences (ICCOINS).* https://doi.org/10.1109/iccoins49721.2021.9497167

Pacton, S., Foulin, J. N., Casalis, S., & Treiman, R. (2013). Children benefit from morphological relatedness when they learn to spell new words. Frontiers in Psychology, 4. https://doi.org/10.3389/fpsyg.2013.00696

Ramli, I., Jamil, N., Seman, N., & Ardi, N. (2015). An improved syllabification for a better Malay language text-to-speech synthesis (TTS). Procedia Computer Science, 76, 417–424. https://doi.org/10.1016/j.procs.2015.12.280

Rosdi, F., Mustafa, M. B., & Salim, S. S. (2017). Assessing automatic speech recognition in measuring speech intelligibility: A study of Malay speakers with speech impairments. *2017 6th International Conference on Electrical Engineering and Informatics (ICEEI).* https://doi.org/10.1109/iceei.2017.8312396

Tan, T.-P., Xiao, X., Tang, E. K., Chng, E. S., & Li, H. (2009). Mass: A Malay language LVCSR Corpus Resource. *2009 Oriental COCOSDA International Conference on Speech Database and Assessments.* https://doi.org/10.1109/icsda.2009.5278382

Tan, T.-P., Goh, S.-S., & Khaw, Y.-M. (2012). A Malay dialect translation and synthesis system: Proposal and preliminary system. 2012 International Conference on Asian Language Processing. https://doi.org/10.1109/ialp.2012.14

Zakaria, N. A., Jamaluddin, S. A., Aslynn, W. A. W., O'Beirne, G., Awatif. N., & Zawani. A. (2021). Development of Pediatric Malay Matrix Sentence Test (PeadMalayMST) materials through construction of Malay corpus.

ORIGINAL ARTICLE

Case Reports of Middle Ear Muscular Spasms Presenting as Tinnitus

Nurlin Binti Ali Hanafiah[1,3,*], Saiful Adli Bin Jamaluddin[1,3], Nur Airin Zulaikha Binti Zamzuri[1], Yahia F. Hussein Al-Hadeethi[2,3], Wan Najibah Wan Mohamad[4]

[1] Department of Audiology and Speech Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[2]Department of Otolaryngology-Head & Neck Surgery, Kulliyyah of Medicine, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia.
[3]Department of Otolaryngology, Sultan Ahmad Shah Medical Centre@IIUM, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[4]Audiology Programme, School of Health Sciences, Universiti Sains Malaysia,16150 Kubang Kerian, Kelantan, Malaysia.

ABSTRACT

Background: Middle ear myoclonus (MEM) is an uncommon disorder characterised by involuntary spasms of the middle ear muscles, often manifesting as objective or pulsatile tinnitus. There are currently no standardised diagnosis or treatment methods, which may result in a chronic impairment in patients' quality of life despite intervention. This study seeks to characterize cases of MEM presented at a tertiary care clinic, with a focus on the diagnostic and management approaches employed. **Methods:** Two clinical cases of MEM were identified from the tinnitus clinic at the Department of Otorhinolaryngology - Head and Neck Surgery, SASMEC @IIUM. Cases were included if they involved patients who had been diagnosed with MEM-induced tinnitus and provided information on diagnostic processes and methods of treatment. **Results:** Two cases of MEM with either objective or subjective tinnitus were presented. Otoscopy and tympanometry were normal across cases, but differences emerged in the use of otoacoustic emissions (OAE), stapedial reflex testing, and pure tone audiometry for assessment. Some cases also utilized imaging to exclude other possible causes. Management strategies ranged from reassurance and pharmacological therapy to surgical tendon sectioning. The SASMEC@IIUM cases favoured conservative approaches: mainly reassurance, counselling, and ongoing monitoring, compared to the more invasive interventions reported in other institutions. **Conclusion:** The observed variability in the diagnosis and management of MEM underscores the lack of standardized clinical guidelines. To facilitate the development of evidence-based protocols, further research involving larger sample sizes and appropriately matched control groups is essential. Such efforts are likely to enhance diagnostic consistency and improve treatment outcomes for patients with MEM-induced tinnitus.

Keywords:
Middle ear myoclonus; tinnitus; audiological evaluation; audiological management

INTRODUCTION

Middle ear muscular spasm, also known as middle ear myoclonus (MEM), is an uncommon yet impactful condition that can lead to objective tinnitus, in which both the patient and examiner can hear the sound (Golz et al., 2003). Swain (2022) categorizes tinnitus, the perception of sound heard without an outside stimulus, as either subjective or objective. Objective tinnitus is far less common and typically results from mechanical or muscle activity within the ear, while subjective tinnitus is more prevalent and associated with hearing loss (Bonaventurová et al., 2022; Golz et al., 2003). MEM-induced tinnitus specifically stems from involuntary contractions of muscles in the middle ear, notably the stapedius and tensor tympani (Ellenstein et al., 2013). Bonaventurová et al (2022) pointed out MEM-induced tinnitus is rare, accounting for only about 1.5% of newly diagnosed tinnitus cases.

This case report aims to explore the impact of MEM-triggered tinnitus on a patient's daily life, while also

* *Corresponding author.*
E-mail address: nurlin@iium.edu.my

discussing the alignment of diagnostic and management approaches used at our centre with those reported in the literature. By examining both clinical outcomes and patient experience, the report seeks to improve understanding of MEM-induced tinnitus and suggest practical strategies for reducing its effect on quality of life.

MATERIALS AND METHODS

This study utilized a multiple clinical case report design, which facilitates the identification of shared and unique diagnostic and treatment features among patients, making it appropriate for reviewing retrospective data while also accounting for the most recent clinical outcomes.

Patient files from the Tinnitus Clinic, Department of Otorhinolaryngology - Head and Neck Surgery, SASMEC @IIUM (from now on referred to as Department of ORL-HNS) were screened to identify potential cases of MEM. Cases were reviewed with an Ear, Nose, and Throat (ENT) specialist where possible MEM cases were selected for each selected case, the patient's medical history, presenting symptoms, and clinical findings were documented. Challenges encountered in diagnosing MEM were also noted during the review process.

Each case was reviewed for audiological assessments such as pure tone audiometry (PTA) and tympanometry. In addition, tinnitus severity and psychological status were evaluated using the *Borang Evaluasi Selidik Tinnitus* (BEST) (Normani et al., 2017) and the Malay version of the Depression Anxiety Stress Scales-21 (DASS-21) (Musa et al., 2007), respectively. Information on treatments, medications, surgery, and alternative options was documented as well. Any other medical conditions that could be related to MEM or tinnitus were also considered. All of this was then discussed with the supervisory team to identify similarities and differences in symptoms, clinical assessment and management, and highlight aspects of good practice. This study was approved by the IIUM Research and Ethics Committee (IREC 2025-KAHS/DASLP8).

RESULTS

Case 1

Case 1 describes a case of 39-year-old Malay woman from the Department ORL-HNS, SASMEC@IIUM who developed left-sided tinnitus five years prior to the first appointment, following the birth of her third child. The tinnitus was triggered by sudden loud sounds, such as her children screaming in enclosed spaces. She also experienced hyperacusis on the affected side. She has been diagnosed with Graves' disease, with a strong family history involving her father and two siblings.

She came to the ENT clinic with complaints of worsening, intermittent left-sided tinnitus. The tinnitus episodes were brief and happened in response to loud acoustic stimulation (hearing her infant cry in a small, enclosed environment). She noted that the tinnitus symptoms began after a brief episode of vertigo and otalgia, which had spontaneously resolved. The vertigo, which was accompanied by brief bouts of imbalance, was produced by postural alterations. During the episode of brief vertigo, she denied experiencing accompanying symptoms such as fever, hearing loss, or ear discharge, but did report prolonged headphone use.

Clinical examination revealed normal otoscopic findings, with intact tympanic membranes. Nasal endoscopy showed minimal mucus accumulation in the nasopharynx. As shown in Figure 1, audiological testing confirmed normal pure tone hearing thresholds and Type A tympanograms in both ears, indicating normal middle ear function.

The initial management was by the ENT specialist which involved prescription of neurotropic vitamins and promoting continuous Valsalva manoeuvres which offered temporary relief. During a session with the audiologist that follows, she described the tinnitus as a soft buzzing and humming sound, triggered by loud noises but reduced in quiet environments. Her tinnitus was graded as very mild on BEST questionnaire, while Case 1's DASS-21 scores showed normal levels of depression and anxiety and mild level of stress. An impression of possible middle ear myoclonus was made by the audiologist, and she was counselled by both the audiologist and the ENT specialist.

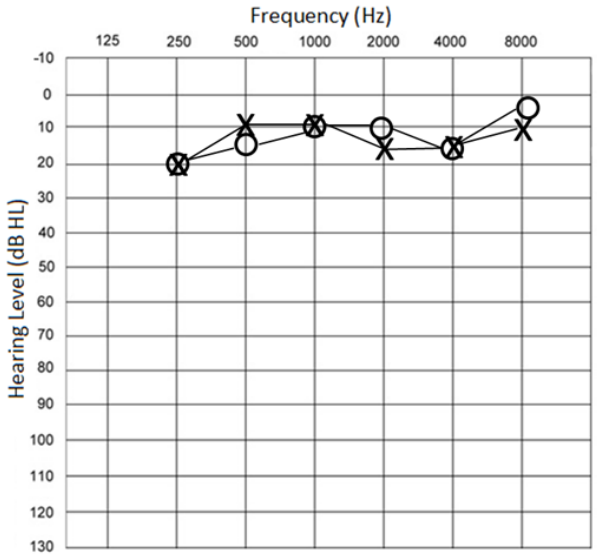


Figure 1. Audiogram of Case 1 showing normal hearing thresholds bilaterally.

Case 2

Case 2 details a 41-year-old Malay man who reported developing right-sided tinnitus after being exposed to a loud insect sound at home. He described the tinnitus as intermittent crackling noise in the right ear that became more noticeable over time, particularly in response to presence of high-pitched sounds. The episodes were also accompanied by headaches and mild giddiness. He denied any associated vertigo, facial palsy, blurred vision, ear discharge, or hearing loss. His medical history included cleft palate repair during infancy.

Figure 2 illustrates the pure tone audiometry findings which revealed essentially normal hearing in both ears. Tympanometry showed Type A tympanograms bilaterally, indicating normal middle ear function. The patient was initially managed by the ENT specialist with a combination of Betahistine hydrochloride and neurotropic vitamins.

Over time, the tinnitus episodes persisted, although they became less frequent and were shorter in duration. He described that the character of the tinnitus later changed to a high-pitched clicking, distorted in quality, and moderate in loudness. The episodes were consistently triggered by specific environmental sounds, including a passing lorry (when he is driving), loud music, his spouse's cough, and tap water splashing on the floor. The symptoms typically resolved shortly after the cessation of the triggering noise and were occasionally relieved by his gentle stroking of hair around his ear.

The patient also observed that the tinnitus became softer during periods of intermittent fasting and worsened upon stopping fasting. He reported a subjective improvement in tinnitus intensity following regular intake of neurotropic vitamins. His BEST score reflected very mild tinnitus severity, and his DASS-21 scores were within normal limits. Based on the clinical presentation and response to sound triggers, an impression of possible MEM in the right ear was made by the audiologist. He was advised to continue taking the neurotropic vitamin supplement and maintain good hydration.

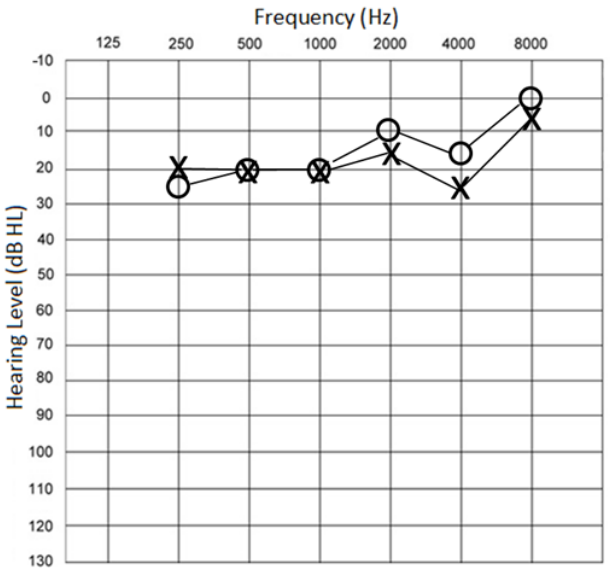


Figure 2. Audiogram of Patient B (Case 2) with essentially normal hearing bilaterally.

Discussion

The present study highlights how the rare condition of MEM can present with recognizable patterns and characteristics. Intermittent, loud-stimulus-induced tinnitus is a consistent feature reported both cases and serves as a powerful diagnostic signal for MEM. Although tinnitus in MEM may be described as buzzing, clicking, or tapping, it often shares mechanical, rhythmic, or percussive features, reflecting its muscular origin (Żyżyńska et al., 2025). From the cases reviewed, both patients experienced intermittent, stimulus-induced tinnitus. The characteristic of the tinnitus varies as it was described as buzzing or humming in Case 1 and crackling in Case 2. The description of clicking is similar to a case

reported by Badia et al. (1994), whereby a 31-year-old woman that was referred from neurological department due to suspicion of palatal myoclonus experienced an audible clicking noise and discomfort in her right ear (from now on referred to as Case 3). On the other hand, Zipfel et al. (2000) described a 44-year-old lady with a history of multiple sclerosis and middle ear myoclonus presented with intermittent "blowing" or "gushing wind" noise bilaterally (from now on referred to as Case 4).

Previous reports suggest symptoms may be worsened by environmental sounds such as loud noises, sudden sound exposure (e.g., an artillery blast, as described by Golz et al., 2003), or physical actions like eye closure or jaw movements (Bhimrao et al., 2012; Ferreira Bento et al., 1998; Ellenstein et al., 2013). This stimulus-triggered pattern is uncommon in typical subjective tinnitus, which is generally constant and internally generated, due to cochlear damage (Swain, 2022; Bonaventurová et al., 2022). Both cases 1 and 2 reported that the tinnitus was triggered by specific loud environmental sounds which is similar with another case reported by Badia et al. (1994) whereby a 28-year-old female reported hearing vibrating or tapping sounds in her left ear which were triggered by specific noises such as phone dial tones, fax machines, bird melodies, and music (from now on referred as Case 5).

Audiological findings revealed that PTA was at normal levels in both cases. This is consistent with Case 3 but differed from the Case 4 that demonstrated bilateral sloping high-frequency sensorineural hearing loss. Normal tympanograms, indicating normal middle function, were also similar features in both Cases 1 and 2. Unfortunately, due to lack of equipment, we are unable to report the middle ear acoustic reflex thresholds which could have provided valuable information regarding the function of the stapedius muscle in response to loud sounds. Case 4 reported rhythmic middle ear movements during tympanometry and physical examination, while Badia et al. (1994) described Case 3 as having normal stapedial reflexes and decay but inconsistent thresholds, suggesting intermittent stapedius involvement. Zipfel et al. (2000) emphasized the diagnostic complexity of MEM and the potential of AR testing. This is supported by Ellenstein et al. (2013) who stated that assessing the brainstem-mediated AR pathway may offer diagnostic insights, particularly in stapedius-related MEM.

Noise-related or stress-related triggers were also noted. In Case 1, tinnitus appeared postpartum and was associated with stress and prolonged headphone usage. Case 2 reported onset after hearing a loud cricket noise and observed that the duration of tinnitus varied during periods of intermittent fasting. Similarly, Case 5 described symptom aggravation linked to temporomandibular joint (TMJ) discomfort and personal stress.

Medical interventions were implemented in most cases. Neurotropic vitamins were prescribed in Cases 1 and 2. Additionally Case 2 also received Betahistine hydrochloride due to previous history of vertigo. In Case 3, benzodiazepines were administered, although they did not provide symptom relief. Surgical intervention, specifically tympanotomy with tendon sectioning, was performed in Cases 3 and 4. Additionally, Case 4 underwent ventilation tube insertion as an early attempt to manage the tinnitus, but it was unsuccessful. Non-pharmacological approaches were also explored. In Case 1, the Valsalva manoeuvre was recommended, while Case 5 focused on stress reduction strategies and received strong reassurance.

Divergent clinical outcomes were noted across the cases under review. In Case 1, the patient reported transient symptomatic relief following the application of the Valsalva manoeuvre. In Case 2, improvements in symptomatology were associated with the administration of neurotropic vitamin supplementation in conjunction with intermittent fasting. Notably, both individuals in Cases 1 and 2 presented with mild tinnitus severity, with Case 1 showing a mild level of stress based on DASS-21 and Case 2 demonstrated no psychological distress. Importantly, the inclusion of the psychological status provided additional insight into the psychosocial dimension of MEM, which was not systematically measured in earlier reports. Case 5, for example, strong reassurance strong reassurance helped reduce the psychological impact of the tinnitus, leading to a more manageable experience. These observations suggest that, in these particular instances, MEM may exert minimal influence on overall quality of life. In comparison, post-surgical relief was achieved in Cases 3 and 4, where both patients reported complete resolution of their tinnitus. Meanwhile, in Case 5, strong reassurance helped reduce the psychological impact of the tinnitus, leading to a more manageable experience.

CONCLUSION

MEM is a rare and frequently underdiagnosed condition, often obscured by its symptomatic overlap with other forms of tinnitus and typically normal audiological findings. However, its hallmark feature- stimulus-dependent, intermittent, and myogenic tinnitus- offer valuable diagnostic cues. MEM is multifactorial in origin, with neurological, psychological, structural, and hormonal influences necessitating comprehensive history taking and a patient-centered management approach. The two cases reported here add to the literature by incorporating standardized psychological assessments, which revealed minimal emotional burden despite persistent symptoms. This dimension, often overlooked in earlier reports, underscores the importance of evaluating both the clinical and psychosocial aspects of MEM. While conservative treatments such as pharmacotherapy and psychological support may suffice for non-objective cases, surgical intervention may be required for persistent or objective tinnitus to achieve lasting resolution. Accurate recognition of MEM's diverse presentations is essential to prevent misdiagnosis and ensure appropriate care. As clinical awareness grows, future research should focus on refining diagnostic criteria that include otological and audiological symptoms and establishing evidence-based treatment protocols that address both otological and psychological impact of MEM.

ACKNOWLEDGEMENT

This research was not funded by any grants. Publication consent was obtained from the Department of Education and Research, SASMEC@IIUM.

REFERENCES

Badia, L., Parikh, A., & Brookes, G. B. (1994). Management of middle ear myoclonus. In *The Journal of Laryngology and Otology* (Vol. 108).

Bhimrao, S. K., Masterson, L., & Baguley, D. (2012). Systematic review of management strategies for middle ear myoclonus. In *Otolaryngology - Head and Neck Surgery (United States)* (Vol. 146, Issue 5, pp. 698–706). https://doi.org/10.1177/0194599811434504

Bonaventurová, M., Bandúrová, V., Čada, Z., Plzák, J., & Balatková, Z. (2022). Middle ear myoclonus as a cause of objective tinnitus. In *Ceska a Slovenska Neurologie a Neurochirurgie* (Vol. 85, Issue 6, pp. 506–508). Czech Medical Association J.E. Purkyne. https://doi.org/10.48095/cccsnn2022506

Ellenstein, A., Yusuf, N., & Hallett, M. (2013). *Reviews Middle Ear Myoclonus: Two Informative Cases and a Systematic Discussion of Myogenic Tinnitus*. http://www.tremorjournal.org

Ferreira Bento, R., Ganz Sanchez, T., Miniti, A., archesi, T.-M., & Paulo, S. (1998). *Continuous, High-frequency Objective Tinnitus Caused by Middle Ear Myoclonus*.

Golz, A., Fradis, M., Netzer, A., Ridder, G. J., Thomas Westerman, S., & Joachims, H. Z. (2003). Bilateral Tinnitus due to Middle-Ear Myoclonus. In *International Tinnitus Journal* (Vol. 9, Issue I).

Musa, R., Fadzil, M. A., & Zain, Z. A. I. N. I. (2007). Translation, validation and psychometric properties of Bahasa Malaysia version of the Depression Anxiety and Stress Scales (DASS). ASEAN Journal of Psychiatry, 8(2), 82-89.

Normani, Z. M., Lau, Y. J., Wan, H. W. S., Lih, A. C., Nik, O. N. A., Salim, R., ... & Wan, M. W. N. (2017). Development and psychometric validation of a new tinnitus questionnaire for clinical use. Auditory & Vestibular Research. 26(2):71-85

Sliwinska-kowalska, M. (2015). Chapter 19 - Hearing. In *Occupational Neurology* (Vol. 131).

Swain, S. K. (2022). Middle ear myoclonus. *BLDE University Journal of Health Sciences*, *7*(2), 179–183. https://doi.org/10.4103/bjhs.bjhs_137_21

Zipfel, T. E., Kaza, S. R., & Greene, J. S. (2000). Clinical Records Middle-ear myoclonus. In *The Journal of Laryngology and Otology* (Vol. 114).

Żyżyńska, K., Grycner, K., Tretiakow, D., & Skorek, A. (2025). Middle ear myoclonus syndrome–a literature review. Polski Przegląd Otorynolaryngologiczny, 14(1), 1-5.

ORIGINAL ARTICLE

Evaluation of Indoor Bacterial Load and Its Association with Respiratory Symptoms among Children in Kuantan Early Childhood Centres

Nur Aliya Maisarah Mohammad Maliki[1], Nurul Falah Mohamad Zaky[1], Noor Fatihah Mohamad Fandi [1*], Norsyuhada Alias[1], Maryam Zahaba[2], Norashikin Sopian[3]

[1]Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia.
[2]Department of Chemistry, Kulliyyah of Science, International Islamic University Malaysia, Pahang, Malaysia.
[3]Faculty of Science and Industrial Technology, University Malaysia Pahang, Malaysia.

Keywords:
Indoor air quality; bioaerosol; respiratory symptoms; children

ABSTRACT

Background: Indoor air quality (IAQ) plays a crucial role in ensuring a healthy and comfortable environment, particularly in early childhood education centers (ECEs). Children are especially vulnerable to poor IAQ due to their developing immune and respiratory systems. This study investigated the indoor bacterial load and the associated respiratory symptoms among children in selected ECEs in Kuantan, Pahang. **Methods:** A cross-sectional study was conducted in two urban ECEs from March to April 2025 to assess indoor bacterial counts and physical parameters (temperature and relative humidity), which were measured with the Surface Air System SAS IAQ 100 and the Testo Air Meter 440, respectively. Potential respiratory symptoms (cough, wheezing, phlegm, and breathlessness) were assessed in 30 children aged from 3 to 6 years using a parent-administered questionnaire adopted from the American Thoracic Society. **Results:** Both ECEs A and B recorded total bacterial counts exceeding the Industry Code of Practice (ICOP) limit, with concentrations of 617.22 CFU/m^3 and 647.22 CFU/m^3, respectively. The recorded temperatures (27.41℃ and 27.96℃) also surpassed the ICOP comfort range. Among the children, cough was the most commonly reported symptom (23.3%), followed by wheezing (16.7%), and phlegm (10.0%). No case of breathlessness symptom was reported in children. Statistical analysis indicated no significant association between the indoor bacterial load and the respiratory symptoms observed. The study was limited by the small sample size and the short monitoring duration, which may not fully represent variations in bacterial levels across different times or seasons. Additionally, the high bacterial counts observed might have been influenced by occupant density, cleaning practices, and ventilation patterns, which were not controlled in this study. However, the present findings provide valuable baseline data on indoor bacterial concentrations, where the local evidence remains limited. **Conclusion:** Overall, the findings indicate early signs that IAQ in ECEs may fall below recommended standards, underscoring the need for better cleanliness and ventilation practices to improve indoor environments and protect children's health. These results also provide useful direction for future research, monitoring efforts, and targeted interventions.

INTRODUCTION

Indoor air quality (IAQ) plays a crucial role in human health, as individuals spend almost 90% of their time indoors (Khamal et al., 2019). The World Health Organization (WHO) reported that indoor air pollution (IAP) is responsible for 3.8 million deaths annually, including 570,000 children under five who die from respiratory infections linked to air pollution and second-hand smoke (WHO, 2017). Monitoring IAQ is therefore essential in early childhood education centres (ECEs), where children are very young, typically between one and six years old, and highly vulnerable due to their immature respiratory and immune systems (Ezinwanne & Okoro, 2023). In addition, children in ECE spend long hours in enclosed indoor environments with high occupant density and having frequent close interactions. Moreover, their behavioural patterns, such as close-contact play, frequent touching of surfaces, and floor-level activities, further increase their susceptibility to airborne bacteria and other indoor biological contaminants.

Indoor pollutants originate from multiple sources, including both indoor and outdoor environments. Physical contaminants such as particulate matter ($PM_{2.5}$ and PM_{10}), dust, and smoke can accumulate in enclosed spaces with inadequate ventilation, aggravating respiratory conditions in children (Pumkaeo et al., 2020). Biological contaminants, including bacteria, fungi, viruses, pollen, and dust mites, are equally significant (US EPA, 2024). Studies in Malaysian schools have reported *Staphylococcus spp.*, *Pseudomonas spp.*, and *Bacillus spp.* as dominant bacterial genera, while *Aspergillus*, *Penicillium*, and *Fusarium* were among the common fungi (Hussin et al., 2011). The presence of these biological pollutants indoors can significantly affect occupants'

*Corresponding author.
E-mail address: fatihahfandi@iium.edu.my

health by causing allergic reactions, respiratory problems, and infections, particularly among children whose immune and respiratory systems are still developing (Ezinwanne & Okoro, 2023; Kumar et al., 2021; EPA, 2024). Continuous exposure to bacteria and fungi in poorly ventilated environments may lead to symptoms such as coughing, wheezing, and phlegm, which could interfere with children's learning and overall well-being (Kumar et al., 2021; Gorny, 2019). Human activities such as coughing, sneezing, and cleaning, as well as the presence of damp or moldy spaces, further contribute to elevated microbial loads indoors (Kumar et al., 2021; Jabeen, 2023).

The Industry Code of Practice on Indoor Air Quality (ICOP IAQ, 2010) provides reference limits for indoor workplace spaces, specifying acceptable ranges for temperature (23-26°C), relative humidity (40-70%), air movement (0.15-0.50 m/s), and biological thresholds of 500 CFU/m^3 for bacteria and 1000 CFU/m^3 for fungi (DOSH, 2010). However, compliance with these standards in indoor playschools and early education settings has not been well documented. Exposure to biological aerosols can cause allergic rhinitis, hypersensitivity, pneumonitis, and communicable diseases such as measles, influenza, and chickenpox (Kumar et al., 2021). Symptoms such as coughing, wheezing, shortness of breath, and lethargy may impair children's well-being and ability to learn (EPA, 2024; Ezinwanne & Okoro, 2023).

Despite the recognised risks, research on IAQ in playschools, particularly among children aged one to six years, remained limited. Prior studies in Kuantan and Pekan either did not report age groups (Hizzri et al., 2018), focused on older children (Hazrin et al., 2018; Madureira et al., 2018), or examined only particulate matter (Maryam et al., 2018). These areas were selected because they represent rapidly developing urban and industrial regions in Pahang, where increased human activity and traffic emissions may influence indoor and outdoor air quality (Nazri et al., 2017; Maryam et al., 2018). Considering that children have higher respiratory rates compared to adults (Choupiliadis & Bhardwaj, 2022), they are more susceptible to bioaerosol exposure. However, empirical evidence linking microbial contamination to health outcomes in this age group is still scarce. This study aims to address this knowledge gap by assessing the IAQ of selected early childhood education centres in Kuantan, Pahang and examining the association between indoor bacterial load and respiratory symptoms among children at ECEs.

MATERIALS AND METHODS

Study Design and Study Location

This cross-sectional study was conducted at two private ECEs in Kuantan, Pahang, Malaysia, referred to as ECE A and ECE B. ECE A is located in Bandar Indera Mahkota, a suburban area of Kuantan that has undergone urban development and mixed residential–commercial land use. Meanwhile, ECE B is located in a mixed commercial–residential area along Jalan Tun Ismail, a major arterial road of Kuantan known for high traffic density and busy commercial activity (Phoo Ken Seng, 2011; Vun Hon Yuan, 2016). The selection of the study sites (Figure 1) was based on accessibility and the centres' willingness to participate.

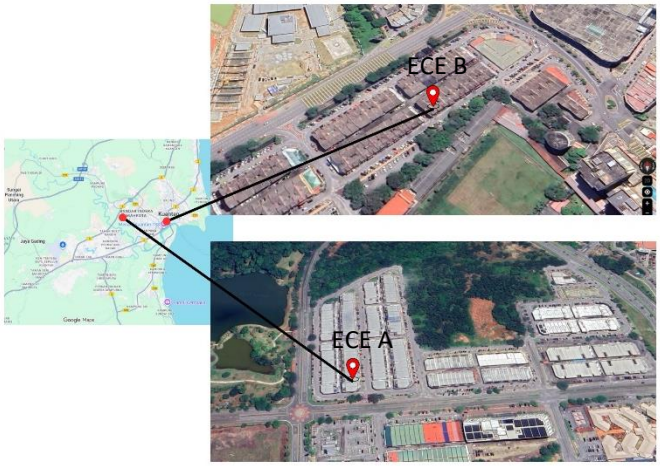


Figure 1: Location of the studied ECE in Kuantan, Pahang. ECE A is located in Bandar Indera Mahkota, while ECE B is situated along Jalan Tun Ismail.

Subject Selection

The study population comprised children aged 3 to 6 years who were enrolled in the selected ECEs in Kuantan. Eligible participants met the following inclusion criteria: children who had attended the ECEs for at least three months to ensure adequate exposure to the ECE's indoor environment, attended the ECE for a full eight-hour day, had no diagnosed chronic respiratory or immunocompromising conditions, and whose parents or guardians provided written informed consent for participation. The inclusion criteria for parents or guardians required that they be the legal parent or guardian of the participating child, be able to comprehend and respond in either Bahasa Melayu or English, be willing to participate in a health-related questionnaire, and reside in Kuantan, living in the same household as the child.

Walkthrough Survey and Questionnaire

A walkthrough survey was conducted to observe the overall conditions of the ECEs, identify suitable sampling locations, and document the types of air-conditioning systems used. Information on children's respiratory symptoms was obtained using a translated version of the American Thoracic Society (ATS-DLD-78-A) questionnaire, originally developed by the American Thoracic Society (Helsing et al., 1979; Smith et al., 2008). The questionnaire also included sections on socio-demographic background and potential home exposure factors. This instrument has been previously validated in local studies conducted by Juliana et al. (2001, 2014) and Sopian et al. (2021). Data were analysed using descriptive statistics to report the prevalence of respiratory symptoms, while Fisher's exact analysis was performed to examine associations between total bacterial count and respiratory health outcomes. All analyses were conducted using IBM Statistical Package for the Social Sciences (SPSS) Version 20.

Indoor Air Quality Assessment

After measuring the floor areas of both ECEs, the total measurement area for each centre was calculated. The number of sampling points was then determined by referring to the Industry Code of Practice on Indoor Air Quality (ICOP, 2010). According to the ICOP, at least one real-time air sampling point is recommended for every 500 m². The measurement area of each ECE was less than 300 m², which meets the requirement for at least one real-time air sampling point per centre. In each ECE, the main activity area and dining area were chosen as sampling sites, as these spaces were commonly used for teaching, learning, and eating activities. During the air sampling, the ECEs were maintained under their normal operating conditions, with air-conditioning units running and all windows kept closed, to accurately reflect the typical indoor environment of the centres.

Testo Air Meter 440 was used to measure physical parameters (temperature and relative humidity). Surface Air System (SAS) IAQ 100 air sampler was used to collect airborne microorganisms by using an agar plate, allowing quantitative total bacterial count analysis. Both instruments were placed at a height of 65-80 cm above the ground (children's breathing zone) and positioned at the main activity area, at least 0.5 m away from walls, corners, and windows. They were oriented away from air-conditioning units and were kept out of direct sunlight, in accordance with ICOP (2010) guidelines. For the SAS IAQ 100 air sampler, the impaction method was employed, whereby air was drawn through the perforated lid at a controlled rate of 100 L/min, thus directing airborne microorganisms onto the surface of the agar plate. After sampling, the agar plate is incubated to allow microbial colonies to grow, and the results are expressed as colony-forming units per cubic meter (CFU/m^3) using the formula shown below (Anua, Haris and Mazlan, 2020):

CFU/m^3 = $[T \times 1000) / [t(min) \times F(L/min)]$

Where:

T: Number of colonies counted from the agar plate

1000: Conversion of a factor of litres to cubic meters

t: Duration of sampling (in minutes)

F: Pump flow rate (L/min)

For physical parameters, data were recorded at 9:00 a.m. to 10:00 a.m., 10:00 a.m. to 11:00 a.m., and 11:00 a.m. to 12:00 p.m. Within each hour, six readings were taken with a 1-minute interval. From this, the average value was determined. Meanwhile, for bacterial load in the indoor ambient, the triplicate technique was used with the device, with 2-minute intervals set for each sampling. The sampling procedure utilised approximately 12 plates between 9.00 a.m. and 12 p.m., with 4 plates used each hour. This IAQ sampling was conducted in March and April 2025, during the normal school operation period. This IAQ sampling was conducted on two separate days within the same week. These devices were only handled by the researcher teams and were not applied to children in schools. Therefore, the safety of children was maintained.

Microbiological Analysis

Nutrient agar was used for environmental monitoring of bacterial sampling. After the data collection, each plate was incubated for 24 to 48 hours at 37°C. Upon the incubation period, the visible colonies of bacteria were counted to determine the total bacterial load. With adjustments made according to the correction table from SAS, the result was expressed in colony-forming units per cubic meter (CFU/m^3).

Ethical Approval

This study was conducted upon the approval of the IIUM Research Ethics Committee (IREC), as it involves human participants with the approval number IREC 2025-KAHS/DBMS15. The ECE staff, parents, and guardians were informed firsthand by explaining the procedure of this study. The permission letter was distributed to parents along with the questionnaire by the teachers. Instructions to use the device were followed by the researcher and

data collection remained anonymous.

RESULTS

Sociodemographic Data of Respondents

A total of 33 parents from both ECEs were approached, but only 30 met the eligibility criteria and participated in the study. Among these 30 children, 17 (56.7%) were male and 13 (43.3%) were female, and all were Malay. More than half of the participants, 17 (56.7%), were enrolled in ECE A, while the remaining 13 (43.3%) were from ECE B. In terms of age distribution, children aged four made up the largest group (36.7%), followed by those aged five (30.0%), six (23.3%), and three (10.0%). No students aged one or two years were enrolled in either of the two ECEs. Most children had been enrolled in their respective ECE for three months to one year (43.4%), while others had been enrolled for two years (10.0%) and three years (3.3%).

Comparison of Building Characteristics of the Studied ECEs

The floor area of ECE A (134 m²) is smaller than that of ECE B (234 m²). Both centres have natural ventilation through windows and mechanical ventilation via air-conditioning; however, only the air-conditioning systems were used in both ECEs. ECE A is equipped with a cassette-type air conditioner, while ECE B uses a split-type air conditioner. During sampling, the air-conditioning units were switched on, and the fans were turned off. Walkthrough inspections revealed that although windows were installed in both ECEs, they remained closed while the air-conditioning was operating. Overall, the general cleanliness in both centres was well maintained. Additionally, both ECEs had a toilet located near the main activity and dining area. Both centres are located in concrete, double-storey shop-lot buildings, with ECE A operating on the upper floor and ECE B on the ground floor.

Total Bacterial Count and Physical Parameters at the Selected ECEs

Table 1 shows that ECE B recorded a slightly higher total bacterial count than ECE A. The bacterial concentrations at ECE A and B were 617.22 CFU/m³ and 647.22 CFU/m³, respectively, corresponding to 1.2 and 1.3 times above the ICOP standard limit (DOSH, 2010). Despite these differences, the total bacterial counts between the two ECEs were not statistically significant (p>0.05). Meanwhile, the measured temperature for both ECEs exceeded the ICOP (DOSH, 2010) guideline, with the mean values of 27.41 and 27.96, respectively. Relative humidity levels were within the acceptable range, recorded at 65.62% for ECE A and 69.38% for ECE B. However, both values were close to the upper permissible limit of 70%.

Table 1: Total Bacterial Count (CFU/m³) and level of Physical Parameter at the Studied ECEs in Kuantan, Pahang.

IAQ Parameter	ECE A Mean ± SD	ECE B Mean ± SD	DOSH ICOP (2010)
Total Bacterial Count (CFU/m³)	**617.22 ± 352.23***	**647.22 ± 338.09***	≤500
Temperature (°C)	**27.41 ± 1.51***	**27.96 ± 1.53***	23-26
Relative Humidity (%)	65.62 ± 9.33	69.38 ± 5.43	40-70

(*) Above the ICOP acceptable limit

Prevalence of Respiratory Symptoms Experienced by Children

Table 2 shows that the respiratory symptoms observed in children at the two ECEs included cough, phlegm, and wheezing, with no case of breathlessness documented. The prevalence of cough was comparable between the centres, at 23.5% in ECE A and 23.1% in ECE B. Phlegm was reported in 11.8% of children in ECE A and 7.7% in ECE B, giving an overall prevalence of 10.0%. Similarly, wheezing was observed in 17.6% of children in ECE A and 15.4% in ECE B, with an overall prevalence of 16.7%. In the overall population of studied children, cough (23.3%) was the most commonly reported symptom, followed by wheezing (16.7%), and lastly phlegm (10.0%).

Table 2: Prevalence of Respiratory Symptoms Experienced by Children ECEs in Kuantan, Pahang.

Respiratory Symptoms	ECE A N = 17(%)	ECE B N = 13(%)	Total N= 30(%)
Cough			
Yes	4 (23.5%)	3 (23.1%)	7 (23.3%)
No	13 (76.5%)	10 (76.9%)	23 (76.7%)
Phlegm			
Yes	2 (11.8%)	1 (7.7%)	3 (10.0%)
No	15 (88.2%)	12 (92.3%)	24 (90.0%)
Wheezing			
Yes	3 (17.6%)	2 (15.4%)	5 (16.7%)
No	14 (82.4%)	11 (84.6%)	25 (83.3%)
Breathlessness			
Yes	0 (0.0%)	0 (0.0%)	0 (0.0%)
No	17 (100.0%)	13 (100.0%)	30 (100.0%)

Association between Total Bacterial Count and

Respiratory Symptoms Experienced by Children in both ECEs

Table 3 presents the association between total bacterial count levels (high > 617.2 CFU/m³ and low ≤ 617.2 CFU/m³) and respiratory symptoms among children in both ECE using Fisher's Exact Test. The analysis showed no statistically significant association between total bacterial count categories and the presence of respiratory symptoms ($p > 0.05$). These findings indicate that, within the sampled ECEs, variations in total airborne bacterial levels did not correspond to differences in reported respiratory symptoms among the children.

Table 3: Association between Total Bacterial Count (CFU/m^3) and Respiratory Symptoms Experienced by Children in both ECEs

Variables	Total Bacteria Count (CFU/m^3)		*p*-value
	High (> 617.2)	Low (≤617.2)	
	n (%)		
Cough			
Yes	4 (57.1)	3 (42.9)	1.000
No	13 (56.5)	10 (43.5)	
Phlegm			
Yes	2 (66.7)	1 (33.3)	1.000
No	15 (55.6)	12 (44.4)	
Wheezing			
Yes	3 (60.0)	2 (40.0)	1.000
No	14 (56)	11 (44.0)	
Breathlessness			
Yes	0 (0)	0 (0)	1.000
No	17 (56.7)	13 (43.3)	

N=30, Fisher's Exact test
(*) Significant at $p < 0.05$

DISCUSSION

Total Bacterial Count in Playschools

In comparing the two ECEs, ECE B demonstrated a higher total bacterial count than ECE A. Although this difference was not statistically significant, it is noteworthy that the total bacterial counts in both centres exceeded the ICOP standard limit by approximately 1.2–1.3 times. The elevated bacterial concentration observed in ECE B could be partly attributed to higher occupancy during sampling, as approximately 35 individuals were present compared with fewer than 30 in ECE A. Hospodsky et al. (2012) reported that human presence can increase airborne bacterial genome concentrations by up to 66-fold in an occupied classroom compared to vacant conditions, contributed largely due to resuspended dust and direct shedding from hair, skin, and nostril passages.

The present findings recorded a slightly higher average temperature in ECE B (~27.96 °C) compared to ECE A (~27.41 °C), which may support enhanced bacterial growth. Moreover, the average relative humidity in ECE B (69.38%) was higher than in ECE A (65.62%), and elevated humidity is known to promote bacterial survival and activity. The combined effect of increased temperature and relative humidity likely contributed to the elevated total bacterial count observed in ECE B, consistent with previous findings showing rapid bacterial growth in hot and humid indoor environments (Albers, 2023; Anualem et al., 2019; Qiu et al., 2022). High relative humidity, in particular, can produce a damp environment that promotes the survival and growth of harmful microorganisms, including bacteria, mold, and viruses (Guarnieri et al., 2023). Additionally, studies in Malaysian laboratories found a positive association between increased indoor temperature and relative humidity and higher bacterial contamination (Yogeswaran et al., 2023).

During the walkthrough inspection, it was observed that although windows were present in both ECEs, none were open, as the indoor spaces were fully air-conditioned. This condition may have indirectly contributed to the accumulation of indoor biological aerosols released by occupants, consistent with findings in a Seremban daycare center, which was also fully air-conditioned (Khamal et al., 2019). Similar observations were reported by Onwusereaka et al. (2022), who found that most schools kept windows closed during school hours, limiting adequate ventilation. Air-conditioned centers generally have lower air-exchange rates compared to naturally ventilated facilities (Khamal et al., 2019), leading to poor ventilation, which further contributes to increased contaminant concentrations and potentially hazardous indoor environments (Albers, 2023). Inadequate ventilation reduces the dilution of pollutants released by occupants, resulting in higher indoor pollutant levels (Guo et al., 2007). Collectively, these findings suggest that the combination of inadequate ventilation, high occupancy, and elevated indoor temperature and humidity likely contributed to the elevated bacterial counts observed in the studied ECEs, which both exceed the ICOP limit of 500

CFU/m^3.

Prevalence of Respiratory Symptoms and Its Association with the Total Count of Bacteria

Based on the present findings, no significant differences in respiratory symptoms were observed between children from the two ECEs, even though total bacterial counts in both centres exceeded recommended limits. Among all participating children, cough (23.3%) was the most commonly reported symptom, followed by wheezing (16.7%) and phlegm (10.0%). These findings are consistent with previous studies in Malaysian early childhood settings. Rawi et al. (2015) reported that preschool children in Selangor showed higher odds of experiencing cough (OR = 3.23; 95% CI: 1.24–8.40) and wheezing (OR = 2.56; 95% CI: 1.02–6.47). Similarly, Khamal et al. (2019) documented a wheezing prevalence of 18.9% among children attending a daycare centre in Seremban.

Although no statistically significant associations between total bacterial counts and respiratory symptoms were detected in this study, airborne bacterial exposure remains a plausible contributor to respiratory issues among young children, as supported by previous literature. Elevated indoor bacterial concentrations are commonly linked to inadequate ventilation and high occupancy density, all of which can irritate the respiratory tract and heighten susceptibility to symptoms (Fisk, 2017; Jaben, 2023; Yusof et al., 2025). Furthermore, routine activities performed by children during school hours, such as breathing, talking, laughing, and sneezing, can release and resuspend biological aerosols into the indoor environment, further increasing airborne microbial load (Hospodsky et al., 2012; Pumkaeo & Iwahashi, 2020).

The insignificant association between total bacterial count and reported respiratory symptoms may be explained by several methodological and participant-related factors. First, the relatively small sample size and the short environmental sampling duration (3–4 hours) may not adequately reflect the children's actual day-to-day exposure levels. Second, the reliance on parent-administered questionnaires introduces limitations in accurately capturing children's respiratory symptoms, particularly when symptoms are mild and easily overlooked. This challenge is more pronounced in children aged 3–6 years, who often struggle to clearly express or describe their physical discomfort. Lampi et al. (2018) reported that older children (9–12 years) can reliably describe their symptoms at a level comparable to parental reporting; however, for younger children, symptom information depends entirely on parental observation. This limitation may also explain the absence of reported breathlessness in the present study, as younger children may not recognise or communicate such experiences effectively. Collectively, these factors, especially the reliance on parent-reported symptoms and smaller sample size, may have contributed to the non-significant associations observed between bacterial exposure and respiratory outcomes.

CONCLUSION

This study evaluated the relationship between bacterial load and respiratory symptoms among children in selected ECEs in Kuantan. Both ECEs recorded bacterial counts exceeding the ICOP limit, alongside slightly elevated temperatures and relative humidity near the upper acceptable range, which are conducive conditions to bacterial growth. Although symptoms of cough, phlegm, and wheezing were reported among children, no significant associations were observed between total bacterial count and respiratory symptoms, likely due to the small sample size. Maintaining good ventilation and cleanliness remains essential to improve indoor air quality and safeguard children's health. Larger studies across more ECEs are recommended to increase statistical power, while future research should include additional IAQ parameters such as volatile organic compounds (VOCs), particulate matter, and fungal counts, and longer data collection periods.

ACKNOWLEDGEMENT

We would like to express our sincere appreciation to the participating ECE centres, parents, and their children for their cooperation, as well as the laboratories from the Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM), for their support throughout this study. Their contributions and willingness to participate were essential to the successful completion of this research.

REFERENCES

Albers, J. (2023, August 31). *What is indoor air quality (IAQ), and how does it impact you?* ISO-Aire Air Purifiers. https://www.iso-aire.com/blog/indoor-air-quality

American Lung Association. (2018, August 22). *50% of Schools Have Poor Indoor Air Quality: How Does Your Child's School Measure Up?* [Press release]. https://www.lung.org/media/press-releases/50-of-schools-have-poor-5utm

Andualem, Z., Gizaw, Z., Bogale, L., & Dagne, H. (2019). Indoor bacterial load and its correlation to physical indoor air quality parameters in public primary schools. *Multidisciplinary Respiratory Medicine*, *14*(2), 1-7. https://doi.org/10.1186/s40248-018-0167-y

Anua, S. M., Haris, N. F. & Mazlan, N. (2020). Bioaerosol Levels in Indoor Air of Animal House and Hospital Laboratories; A Comparison With Library and Administrative Offices. *Malaysian Journal of Medicine and Health Sciences*, 16(SUPP 1), 94-100. Retrieved at https://medic.upm.edu.my/upload/dokumen/20200102150414 15_MJMHS_0105.pdf

Fisk, W. J. (2017). The ventilation problem in schools: literature review. *Indoor air, 27*(6), 1039-1051. https://doi.org/10.1111/ina.12403

Guo, H., Morawska, L., He, C., & Gilbert, D. (2007). Impact of ventilation scenario on air exchange rates and on indoor particle number concentrations in an air-conditioned classroom. *Atmospheric Environment*, *42*(4), 757–768. https://doi.org/10.1016/j.atmosenv.2007.09.070

Guarnieri, G., Olivieri, B., Senna, G., & Vianello, A. (2023). Relative humidity and its impact on the immune system and infections. *International Journal of Molecular Sciences*, *24*(11), Article 9456. https://doi.org/10.3390/ijms24119456

Hazrin, A. H., Maryam, Z., Hizrri, A., Norhidayah, A., Samsuddin, N., & Mohd Shukri, M. A. (2017). Occupancy Implications on Indoor Air Quality (IAQ) in selected primary school classrooms around Kuantan, Pahang. Malaysian Journal of Public Health Medicine, *Special volume* (1), 95-105. https://umpir.ump.edu.my/id/eprint/21796/

Hospodsky, D., Qian, J., Nazaroff, W. W., Yamamoto, N., Bibby, K., Rismani-Yazdi, H., & Peccia, J. (2012). Human occupancy as a source of indoor airborne bacteria. *PLoS ONE*, *7*(4), Article e34867. https://doi.org/10.1371/journal.pone.0034867

Jabeen, R., Kizhisseri, M. I., Mayanaik, S. N., & Mohamed, M. M. (2023). Bioaerosol assessment in indoor and outdoor environments: A case study from India. *Scientific Reports*, *13*(1), Article 18066. https://doi.org/10.1038/s41598-023-44315-z

Khamal, R., Isa, Z. M., Sutan, R., Noraini, N. M. R., & Ghazi, H. F. (2019). Indoor Particulate Matters, Microbial Count Assessments, and Wheezing Symptoms among Toddlers in Urban Day Care Centres in the District of Seremban, Malaysia. *Annals of Global Health, 85*(1), 1-15. https://doi.org/10.5334/aogh.2425

Lampi, J., Ung-Lanki, S., Santalahti, P., & Pekkanen, J. (2018). Test-retest repeatability of child's respiratory symptoms and perceived indoor air quality – comparing self- and parent-administered questionnaires. *BMC Pulmonary Medicine*, *18*(1), 1-7. https://doi.org/10.1186/s12890-018-0584-x

Madureira, J., Aguiar, L., Pereira, C., Mendes, A., Querido, M. M., Neves, P., & Teixeira, J. P. (2018). Indoor exposure to bioaerosol particles: Levels and implications for inhalation dose rates in schoolchildren. *Air Quality, Atmosphere & Health*, *11*(8), 955–964. https://doi.org/10.1007/s11869-018-0599-8

Maryam, Z., Hazrin, A. H., Hizrri, A., Norhidayah, A., Samsuddin, N., & Mohd Shukri, M. A. (2018). Association of particulate matter (PM) with respiratory symptoms among children in selected primary schools in Pahang. *J Clean WAS*, *2*(1), 11-15, Retrieved at https://jcleanwas.com/archives/1jcleanwas2018/1jcleanwas2018-11-15.pdf

Onwusereaka, C. O., Jalaludin, J., & Hisamuddin, N. H. (2022). Indoor Air Quality and Respiratory Health Implication Among Malay Preschool Children in Puchong and Hulu Langat Selangor, Malaysia. *Malaysian Journal of Medicine & Health Sciences*, *18*(SUPP5), 62-74. Retrieved at http://psasir.upm.edu.my/id/eprint/101877/1/20220411095501 8_1310.pdf

P, N. (2012). Asthma Severity and Environmental Health Risk Factor among Asthmatic Primary School Children in the Selected Areas. *American Journal of Applied Sciences*, *9*(10), 1553–1560. https://doi.org/10.3844/ajassp.2012.1553.1560

Phoo Ken Seng.(2011). *Proposal to improve parking facilities along Jalan Tun Ismail.* [Bachelor's Degree thesis, Universiti Malaysia Pahang]. Open Access Repository of UMPSA Research and Publication. https://umpir.ump.edu.my/id/eprint/7770

Pumkaeo, P., & Iwahashi, H. (2020). Bioaerosol sources, sampling methods, and major categories: A comprehensive overview. *Reviews in Agricultural Science*, *8*(0), 261–278. https://doi.org/10.7831/ras.8.0_261

Qiu, Y., Zhou, Y., Chang, Y., Liang, X., Zhang, H., Lin, X., Qing, K., Zhou, X., & Luo, Z. (2022). The effects of ventilation, humidity, and temperature on bacterial growth and bacterial genera distribution. *International Journal of Environmental Research and Public Health*, *19*(22), Article 15345. https://doi.org/10.3390/ijerph192215345

Rawi, N. a. M. N., Jalaludin, J., & Chua, P. C. (2015). Indoor Air Quality and Respiratory Health among Malay Preschool Children in Selangor. *BioMed Research International*, 2015, 1–8, Article 248178. https://doi.org/10.1155/2015/248178

Sopian, N. A., Jalaludin, J., Abu Bakar, S., Hamedon, T. R., & Latif, M. T. (2021). Exposure to particulate PAHs on potential genotoxicity and cancer risk among school children living near the petrochemical industry. *International Journal of Environmental Research and Public Health*, *18*(5), Article 2575. https://doi.org/10.3390/ijerph18052575

U.S. Environmental Protection Agency. (n.d.). *Section 2: Indoor air quality*. https://www.epa.gov/sites/default/files/2014-08/documents/sec_2.pdf

Vun Hon Yuan. (2016). *Effect Of Implementing One Way Traffic In Kuantan City.* [Bachelor's Degree thesis, Universiti Malaysia Pahang]. Open Access Repository of UMPSA Research and Publication https://umpir.ump.edu.my/id/eprint/16070

World Health Organization. (2017, March 6). *The cost of a polluted environment: 1.7 million child deaths a year, says WHO*. Retrieved at https://www.who.int/news/item/06-03-2017-the-cost-of-a-polluted-environment-1-7-million-child-deaths-a-year-says-who

Yogeswaran, K., Azmi, L., Bhassu, S., Isa, H. N., & Aziz, M. A. (2023). Physical parameters influence the microbial quality of indoor air in research laboratories: A report from Malaysia. *Kuwait Journal of Science*, *50*(4), 665-673. https://doi.org/10.1016/j.kjs.2023.04.009

Yusof, S. K., Abdull, N., Suhaimi, N. S., Abdullah Suhaimi, A., & Moh Hussin, N. (2025). Bioaerosol concentration and health implication for children in daycare centers. *Aerobiologia,* *41*(3), 643-665. https://doi.org/10.1007/s10453-025-09872-y

ORIGINAL ARTICLE

Social Media Addiction and Stress as Predictors of Eating Behaviour Among Undergraduate Students: A Cross-Sectional Study

Hannah Sorfina Azmi[1], Muhamad Ariff Ibrahim[1,*], Mohd Nazir Mohd Nazori[2], Nurulwahida Saad[3] & Siti Adibah Waisulqrnai[3]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2]Department of Physical Rehabilitation Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[3]Department of Biomedical Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

ABSTRACT

Introduction: The increasing influence of social media and stress on young adults has raised concerns regarding their impact on eating behaviour. University students, in particular, are vulnerable to poor coping strategies that may lead to disordered eating patterns, especially when exposed to high stress and excessive social media use. This study aimed to determine the levels of social media addiction, stress, and eating behaviour among undergraduate students at a Malaysian public university, and to examine their associations and predictive relationships. **Methods:** A quantitative cross-sectional study was conducted among 383 undergraduate students from all campuses using validated instruments: the Social Media Addiction Scale – Student Form (SMAS-SF), the Perceived Stress Scale (PSS), and the Dutch Eating Behaviour Questionnaire (DEBQ). **Results:** Correlation analysis showed that stress and social media dimensions such as virtual tolerance, virtual problem, and virtual information were significantly related to at least one type of eating behaviour dimension. Multiple regression analysis showed that stress ($\beta = 0.103$, $p = 0.048$) was a significant predictor of restrained eating. Virtual problem ($\beta = 0.132$, $p = 0.034$) was a significant predictor of emotional eating behaviour. Additionally, virtual tolerance ($\beta = 0.141$, $p = 0.020$), virtual information ($\beta = 0.114$, $p = 0.042$), and stress ($\beta = -0.156$, $p = 0.002$) were significant predictors of external eating behaviour. **Conclusion:** The study concludes that both social media addiction and stress play a role in shaping students' eating behaviours, with different components influencing restrained, emotional, and external eating in varying directions. These findings suggest the need for varied approaches when addressing unhealthy eating patterns among university students.

Keywords:
social media additions, stress, eating behaviour, & university students.

INTRODUCTION

Eating behaviours play an important role in maintaining physical health and improving the quality of life, as well as in preventing chronic diseases. A healthy diet is essential to protect against malnutrition in all its forms and noncommunicable diseases (World Health Organization, 2020).

Young adults, particularly university students, often struggle to maintain a healthy diet as they find themselves in new social settings and making their own food choices for the first time. Among these, stress conditions and social media exposure have emerged as influential factors affecting food choices and eating patterns. University life presents unique stressors, including academic demands, social adjustments, and personal changes, which can significantly impact students' mental health and coping behaviour. Studies have shown that high stress levels are associated with disordered eating, such as emotional overeating or undereating, as

* *Corresponding author.*
E-mail address: ariffib@iium.edu.my

individuals turn to food for comfort or control (Dakanalis et al., 2023; Choi, 2020). At the same time, the widespread use of social media among university students exposes them to idealized body images, viral diet trends, and comparison culture, which can further drive unhealthy eating behaviours and body dissatisfaction (Aparicio-Martinez et al., 2019; Jiotsa et al., 2021).

The combination of stress and social media pressure creates a complex environment where students may engage in restrictive dieting, binge eating, or meal skipping, which are often reinforced by peer influence and the desire for social acceptance (Ragelienė & Grønhøj, 2020). These behaviours not only affect students' physical health but also increase their risk for mental health issues and long-term disordered eating patterns. In Malaysia, rising rates of adolescent mental health concerns and obesity emphasise the urgency of addressing these interrelated challenges (Institute for Public Health, 2017; Shah et al., 2023). In light of these concerns, the present study aims to determine the levels of social media addiction, stress, and eating behaviour among students, and to examine how these factors are associated and influence one another.

MATERIAL AND METHODS

Study Participants

This research employed a quantitative, cross-sectional study to examine the relationships between social media addiction, stress, and eating behaviours among undergraduate students at a public Malaysian university. Data collection was carried out through both online and paper-based surveys, targeting students from three main campuses: Gombak (14,994 students), Kuantan (3,664 students), and Pagoh (1,662 students), with a total student population of 20,320 at the time of the study.

The Kuantan Campus has six faculties: Kulliyyah of Allied Health Sciences, Dentistry, Medicine, Nursing, Pharmacy, and Science. The Gombak campus includes seven faculties: Kulliyyah of Islamic Revealed Knowledge and Human Sciences, Law, Architecture and Environmental Design, Economics and Management Sciences, Education, Engineering, and Information and Communication Technology. Meanwhile, the Pagoh campus hosts a single faculty, the Kulliyyah of Sustainable Tourism and Contemporary Languages. Based on Krejcie & Morgan's (1970) sample size calculation, 379 participants were required for a population of this size. Ultimately, 383 students from all campuses participated in the study, selected through convenience sampling.

Questionnaire

A comprehensive questionnaire was developed using the adopt-and-adapt method, drawing on validated instruments by Ursoniu et al. (2022), Cohen et al. (1994), and Van Strien et al. (1986). The survey was administered online via Google Forms and distributed through email, WhatsApp, Telegram, and Instagram. To maximize reach, paper-based copies were also provided on campus for students who preferred or had better access to hard-copy questionnaires. The survey consisted of four sections: sociodemographic information, social media addiction, stress levels, and eating behaviour.

The survey included six sociodemographic items, namely gender, age, race, year of study, student origin, and Kulliyyah. The Social Media Addiction Scale – Student Form (SMAS-SF), developed by Ursoniu et al. (2022), was used to assess social media addiction among university students. The scale consisted of 29 items across four dimensions (virtual tolerance refers to increasing time spent on social media to achieve satisfaction, similar to the tolerance seen in other addictions. Virtual communication is involves using social media to manage emotions and is often tied to limited self-presentation skills. Virtual problem encompasses issues arising from excessive use, such as reduced productivity, academic struggles, and health effects. Lastly, virtual information reflects continuous searching for updates, leading to information overload and potentially stimulating addiction pathways.) using a 5-point Likert scale. Total scores range from 29 to 145, with higher scores indicating a greater level of social media addiction. The Perceived Stress Scale (PSS), developed by Cohen et al. (1983), was used to measure perceived stress levels. The 10-item version assesses the degree to which situations in life are perceived as stressful. Scores range from 0 to 40, categorized as low (0–13), moderate (14–26), and high (27–40) stress. The Dutch Eating Behaviour Questionnaire (DEBQ), by Van Strien et al. (1986), was used to measure three types of eating behaviour: emotional is triggered by emotional responses; external is a response to environmental cues; and restrained eating is referred to as controlled or limited food intake. The 5-point Likert scale ranges from

"Never" to "Very Often," with higher average scores indicating stronger tendencies in each category.

Statistical Analysis

The Statistical Package for the Social Sciences (SPSS) version 29.0 was used to analyse the data. Descriptive statistics were applied to summarize sociodemographic variables (percentage, mean, SD), social media addiction levels (SMAF-SF), and stress levels (PSS). Additionally, the relationship between eating behaviours, social media addiction, and stress was examined using Pearson Correlation and multiple regression

RESULTS

Demographic

According to Table 4.1, the majority of respondents were female students, accounting for 80.2% (n=307), while male students comprised 19.8% (n=76). Most respondents were aged between 20 and 30 years, representing 90.9% (n=348), followed by those aged 23 to 25 years at 9.1% (n=35). In terms of race, the majority were Malay (98.2%, n=376), with smaller proportions of Chinese (1.0%, n=4), Indian (0.5%, n=2), and an international student categorized under 'Other' (0.3%, n=1).

In terms of the year of study, the majority of respondents were from Year 1, comprising 30.0% (n=115), followed by Year 3 students at 27.7% (n=106). Respondents from Year 2 accounted for 23.2% (n=89), while those from Year 4 made up 19.1% (n=73). Regarding students' origin, as previously mentioned, most were local students (98.4%, n=377), with a small proportion of international students (1.6%, n=6). The majority of respondents were from the Kuantan campus (76.4%), primarily from the Kulliyyah of Allied Health Sciences (KAHS), which accounted for 51.2% (n=196). Meanwhile, 21.8% of respondents were from the Gombak campus, with the highest representation from the Kulliyyah of Architecture and Environmental Design (KAED) at 8.6% (n=33). A smaller portion, 1.8% (n=7), were from the Pagoh campus, which consisted of students from the Kulliyyah of Sustainable Tourism and Contemporary Languages (KSTCL).

Table 1: Sociodemographic Factors of Respondents

Characteristics	Categories	n	%
Gender	Male	76	19.8
	Female	307	80.2
Age	19 – 22	348	90.9
	23 - 25	35	9.1
Race	Malay	376	98.2
	Chinese	4	1.0
	Indian	2	0.5
	Others	1	0.3
Year of study	Year 1	115	30.0
	Year 2	89	23.2
	Year 3	106	27.7
	Year 4	73	19.1
Student origin	Local	377	98.4
	International	6	1.6
Kulliyyah	KAHS	196	51.2
	KOD	12	3.1
	KOP	21	5.5
	KOM	33	8.6
	KOS	20	5.2
	KON	10	2.6
	KICT	12	3.1
	KOED	10	2.6
	KOE	15	3.9
	KIRKHS	7	1.8
	KSTCL	7	1.8
	KENMS	7	1.8
	KAED	33	8.6

Level of Social Media Addiction and Stress

As shown in Table 2, social media addiction levels were categorized into two groups: 'Low' (scores 29–86) and 'High' (scores 87–145). The majority of students (n = 258; 67.4%) fell into the low addiction category, whereas 125 students (32.6%) were classified as having a high level of social media addiction. Meanwhile, the stress scores were grouped into three distinct categories: 'Low' (score range: 0 to 13), 'Moderate' (14 to 26), and 'Perceived High' (27 to 40). These classifications were based on established scoring criteria to better understand the mental well-being of the students. The results indicated that most respondents experienced a moderate level of stress, with 277 students (72.3%) falling into this category. In

comparison, 57 students (14.9%) reported high perceived stress, while 49 students (12.8%) reported low stress levels.

Table 2: Level of Social Media Addiction and Stress

Variables	Level	n	%
Social Media Addiction	Low	258	67.4
	High	125	32.6
Stress	Low	49	12.8
	Moderate	277	72.3
	High	57	14.9

Correlation

Table 3 presents the correlation analysis results examining the relationships between four dimensions of social media addiction (Virtual Tolerance, Virtual Communication, Virtual Problem, and Virtual Information) and stress with three types of eating behaviours (Restrained, Emotional, and External eating) among undergraduate students. The findings indicate that Virtual Tolerance is significantly associated with all three eating behaviours: Restrained ($r = 0.119$, $p < 0.05$), Emotional ($r = 0.131$, $p < 0.05$), and External eating ($r = 0.132$, $p < 0.05$). In contrast, Virtual Communication shows no significant correlation with any of the eating behaviour types, as all p-values exceed 0.05, indicating no meaningful relationship. For Virtual Problem, significant correlations were found with Emotional Eating ($r = 0.161$, $p < 0.05$) and External Eating ($r = 0.201$, $p < 0.05$), while no significant relationship was observed with Restrained Eating ($r = 0.069$, $p > 0.01$). Regarding Virtual Information, significant relationships were identified with Restrained Eating ($r = 0.132$, $p < 0.05$), Emotional Eating ($r = 0.121$, $p < 0.05$), and External Eating ($r = 0.184$, $p < 0.05$). Lastly, for the Stress variable, significant associations were observed with Restrained Eating ($r = 0.106$, $p < 0.05$) and External Eating ($r = 0.131$, $p < 0.05$). No significant relationship was found between stress and Emotional Eating ($r = -0.054$, $p > 0.05$). In summary, the statistically significant relationships observed were accompanied by low r-values, implying weak correlations and the potential influence of other factors on eating behaviours.

Table 3: Correlation analysis between social media addiction, stress and eating behaviour

Variables	1	2	3	4	5	6	7
1. Virtual Tolerance	-						
2. Virtual Communication	.585**	-					
3. Virtual Problem	.529**	.642**	-				
4. Virtual Information	.418**	.331**	.400**	-			
5. Stress	.079	.189**	.190**	.052	-		
6. Restrained Eating	.119*	.066	.069	.132**	.106*	-	
7. Emotional Eating	.131*	.033	.161**	.121*	-.054	-.005	-
8. External Eating	.201*	.019	.136**	.184**	-.131*	-.163**	.381**

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Multiple Regression Analysis

Table 4 summarizes the multiple regression findings on the impact of social media addiction and stress on restrained eating. The overall model was statistically significant, F(4, 378) = 2.546, **p = 0.028**, indicating a meaningful relationship between the predictors and the outcome. The model explained **3.3**% of the variance in restrained eating (R = 0.181, R^2 = 0.033), with an adjusted R^2 of 0.020. Moreover, Table 5 shows that stress was the only significant predictor of restrained eating (B = 0.150, p **= 0.048**), suggesting that higher stress levels were associated with increased restrained eating. Other variables—**virtual tolerance** (p = 0.695), **virtual problem** (p = 0.710), and **virtual information** (p = 0.066)—were not statistically significant.

Table 4: Regression analysis between social media addiction and stress towards restrained eating behaviour

Model	B	SE	Beta	t	Sig.
Virtual Tolerance	.210	.144	.097	1.453	.147
Virtual Communication	-.044	.112	-.028	-.392	.695
Virtual Problem	-.043	.115	-.026	-.373	.710
Virtual Information	.231	.125	.106	1.844	.066
Stress	.150	.075	.103	1.986	.048
Model Summary	***R= .181***				
	R^2 = .033				
	Adj. R^2 = .020				
	SE = 9.896				
	F (4,378) = 2.546				

Table 5 presents the multiple regression analysis examining the effect of social media addiction and stress on emotional eating behaviour. The model was statistically significant, F(4, 378) = 3.762, p = 0.005, indicating that the predictors had a meaningful impact. The model explained 3.8% of the variance in emotional eating (R = 0.196, R^2 = 0.038), with an adjusted R^2 of 0.028.

Among the predictors, only virtual problem showed a significant positive effect (B = 0.254, p = 0.034), suggesting that other issues related to problematic social media use may be linked to higher emotional eating. Other variables, which are virtual tolerance (p = 0.457), virtual information (p = 0.353), and stress (p = 0.099), were not statistically significant.

Table 5: Regression analysis between social media addiction and stress towards emotional eating behaviour

Model	B	SE	Beta	t	Sig.
Virtual Tolerance	.117	.157	.046	.744	.457
Virtual Problem	.254	.120	.132	2.126	.034
Virtual Information	.137	.147	.053	.930	.353
Stress	-.146	.088	-.085	-1.653	.099
Model Summary	***R=.196***				
	R^2 = .038				
	Adj. R^2 = .028				
	SE = 10.430				
	F (4,378) = 3.762				

Table 6 presents the multiple regression analysis on the effect of social media addiction and stress on external eating behaviour. The model was statistically significant, F(4, 378) = 7.771, p < 0.001, indicating a meaningful relationship between the predictors and external eating. The model explained 7.6% of the variance (R = 0.276, R^2 = 0.076), with an adjusted R^2 of 0.066.

Three predictors were statistically significant: virtual tolerance (B = 0.182, p = 0.020) and virtual information (B = 0.149, p = 0.042) were positively associated with external eating, while stress showed a significant negative relationship (B = -0.136, p = 0.002). Virtual problem (p = 0.451) was not a significant predictor.

Table 6: Regression analysis between social media addiction and stress towards external eating behaviour

Model	B	SE	Beta	t	Sig.
Virtual Tolerance	0.182	0.078	0.141	2.329	0.020
Virtual Problem	0.045	0.060	0.046	0.755	0.451
Virtual Information	0.149	0.073	0.114	2.043	0.042
Stress	-	0.044	-	-	0.002

	0.136	0.156	3.104

Model Summary	***R= .276***
	R^2 = .076
	Adj. R^2 = .066
	SE = 5.191
	F (4,378) = 7.771

DISCUSSION

The present study revealed that most students demonstrated low levels of social media addiction, consistent with the findings of Lokman et al. (2023) and Zaiamri & Abidin (2023), who reported that Malaysian students primarily use social media for functional purposes, such as academic and professional activities. Nevertheless, a notable proportion of students demonstrated high levels of addiction, which may adversely affect mental well-being and daily routines, as highlighted by Akakandelwa and Walubita (2017) and Pellegrino et al. (2022).

Most students also reported moderate stress levels, consistent with the findings of Selamat et al. (2024), Wong et al. (2023), and Asif et al. (2020), highlighting that psychological distress is a persistent issue in university settings. Stress remains a critical factor influencing both health and behavioural outcomes, including dietary habits. As a result, several researchers have suggested that food intake following stressful events should align with the body's natural physiological needs (Schultchen et al., 2019; Radzi et al., 2022).

Regarding eating behaviours, restrained eating was significantly associated with higher stress levels, suggesting that stress may lead to restrictive eating patterns as a form of self-control and coping mechanism in response to psychological strain or body image concerns, consistent with the findings of Yong et al. (2021) and Smith et al. (2022). Emotional eating was significantly predicted by virtual problems, a dimension of social media addiction, indicating that difficulty in disengaging from online platforms could contribute to maladaptive coping through food, supporting Rossi (2025) and Paulus et al. (2021). However, stress was not significantly related to emotional eating, which may reflect diverse coping mechanisms such as spiritual or social support often used by the students.

External eating was most strongly linked to both virtual tolerance and virtual information. This suggests that exposure to food content on social media may trigger eating responses based on visual cues rather than hunger, aligning with Morales & Berridge (2020) and Spence et al. (2016). Interestingly, stress had a negative association with external eating, potentially due to attentional narrowing or appetite suppression during academic demands, as supported by Nakamura et al. (2020) and Chu et al. (2019).

Overall, the findings emphasize that social media addiction and stress influence eating behaviours in different ways, highlighting the importance of targeted interventions focusing on digital literacy and stress management among university students.

CONCLUSION

This study explored the relationship between social media addiction, stress, and eating behaviours among undergraduate students. Results showed that while most students experienced moderate stress, many reported low social media addiction. Emotional eating was mainly linked to virtual problems, restrained eating was influenced by stress, and external eating was associated with both virtual tolerance, virtual information, and negatively with stress. These findings suggest that social media and stress affect eating patterns in different ways. Addressing both digital habits and psychological stress is essential in developing targeted interventions to promote healthier eating behaviours among university students.

RECOMMENDATIONS

Based on the findings, it is recommended that universities implement targeted health programs focusing on stress management, digital well-being, and healthy eating behaviours. Increasing awareness among students about the negative impacts of excessive social media use on both mental and physical health is also important. This can be achieved through workshops and social media literacy modules to help reduce harmful influences. Additionally, mental health services within the university should be strengthened, with particular attention given to addressing stress-related eating behaviours. For future research, studies should explore the long-term effects of social media and stress on eating behaviours, consider qualitative or mixed-method approaches, and assess additional moderating factors such as self-esteem, resilience, and social support. It is also encouraged that future studies apply relevant theoretical frameworks to improve the accuracy and comprehensiveness of research related to eating behaviours.

ACKNOWLEGMENT
This research was not funded by any grant.

REFERENCES

Akakandelwa, A., & Walubita, G. (2017). Students' social media use and its perceived impact on their social life: A case study of the University of Zambia. *The International Journal of Multi-Disciplinary Research*, *October*.

Aparicio-Martinez, P., Perea-Moreno, A. J., Martinez-Jimenez, M. P., Redel-Macías, M. D., Pagliari, C., & Vaquero-Abellan, M. (2019). Social media, thin-ideal, body dissatisfaction and disordered eating attitudes: An exploratory analysis. *International Journal of Environmental Research and Public Health*, *16*(21). https://doi.org/10.3390/ijerph16214177

Choi, J. (2020). Impact of stress levels on eating behaviors among college students. *Nutrients*, *12*(5). https://doi.org/10.3390/nu12051241

Chu, B., Marwaha, K., & Ayers, D. (2019). Physiology, Stress Reaction. In *StatPearls*.

Cohen, S. (1994). *PERCEIVED STRESS SCALE*. www.mindgarden.com

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. In *Source: Journal of Health and Social Behavior* (Vol. 24, Issue 4). https://about.jstor.org/terms

Dakanalis, A., Mentzelou, M., Papadopoulou, S. K., Papandreou, D., Spanoudaki, M., Vasios, G. K., Pavlidou, E., Mantzorou, M., & Giaginis, C. (2023). The Association of Emotional Eating with Overweight/Obesity, Depression, Anxiety/Stress, and Dietary Patterns: A Review of the Current Clinical Evidence. In *Nutrients* (Vol. 15, Issue 5). MDPI. https://doi.org/10.3390/nu15051173

Institute for Public Health, M. O. H. M. (2017). *National Health and Morbidity Survey 2017*.

Jiotsa, B., Naccache, B., Duval, M., Rocher, B., & Grall-Bronnec, M. (2021). Social media use and body image disorders: Association between frequency of comparing one's own physical appearance to that of people being followed on social media and body dissatisfaction and drive for thinness. *International Journal of Environmental Research and Public Health*, *18*(6), 1–14. https://doi.org/10.3390/ijerph18062880

Lokman, H., Ibrahim, M. A., Mohd Nazori, M. N., Saad, N., & Sahid, N. I. A. (2023). The Correlation of Anxiety and Social Media Addiction on Eating Behavior Among IIUM Students. *International Journal of Allied Health Sciences*, *Vol. 7*(No. 5).

Morales, I., & Berridge, K. C. (2020). 'Liking' and 'wanting' in eating and food reward: Brain mechanisms and clinical implications. In *Physiology and Behavior* (Vol. 227). https://doi.org/10.1016/j.physbeh.2020.113152

Nakamura, C., Ishii, A., Matsuo, T., Ishida, R., Yamaguchi, T., Takada, K., Uji, M., & Yoshikawa, T. (2020). Neural effects of acute stress on appetite: A magnetoencephalography study. *PLoS ONE*, *15*(1). https://doi.org/10.1371/journal.pone.0228039

Paulus, F. W., Ohmann, S., Möhler, E., Plener, P., & Popow, C. (2021). Emotional Dysregulation in Children and Adolescents With Psychiatric Disorders. A Narrative Review. In *Frontiers in Psychiatry* (Vol. 12). https://doi.org/10.3389/fpsyt.2021.628252

Pellegrino, A., Stasi, A., & Bhatiasevi, V. (2022). Research trends in social media addiction and problematic social media use: A bibliometric analysis. In *Frontiers in Psychiatry* (Vol. 13). https://doi.org/10.3389/fpsyt.2022.1017506

Radzi, K. N., Ibrahim, M. A., Saad, N., Mohd Nazori, M. N., & Shamsuddin, A. S. (2024). View of academic stress and emotional eating behaviour among IIUM students. *International Journal of Allied Health Sciences*. Iium.edu.my. https://journals.iium.edu.my/ijahs/index.php/IJAHS/article/view/795/681

Ragelienė, T., & Grønhøj, A. (2020). The influence of peers' and siblings' on children's and

adolescents' healthy eating behavior. A systematic literature review. *Appetite*, *148*. https://doi.org/10.1016/j.appet.2020.104592

Rossi, A. A. (2025). Tying Food Addiction to Uncontrolled Eating: The Roles of Eating-Related Thoughts and Emotional Eating. *Nutrients* , *17*(3). https://doi.org/10.3390/nu17030369

Schultchen, D., Reichenberger, J., Mittl, T., Weh, T. R., Smyth, J. M., Blechert, J., & Pollatos, O. (2019). Bidirectional relationship of stress and affect with physical activity and healthy eating.British journal of health psychology,24(2), 315-333

Selamat, A. S., Ibrahim, M. A., Nazori, M. N., Saad, N., & Waisulqrnai, S. A. (2024). View of Relationship Between Body Image Perception and Stress Towards Eating Behaviour Among University Students. *2024*, *Vol. 8*(No. 5).

Shah, N. M., Aghamohammadi, N., Thangiah, N., Ng, A. K., & Majid, H. A. (2023). Association between stress and eating behaviour among Malaysian adolescents prior to examination. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-34699-3

Smith, K. E., Mason, T. B., Wang, W. L., Schumacher, L. M., Pellegrini, C. A., Goldschmidt, A. B., & Unick, J. L. (2022). Dynamic associations between anxiety, stress, physical activity, and eating regulation over the course of a behavioral weight loss intervention. *Appetite*, *168*. https://doi.org/10.1016/j.appet.2021.105706

Spence, C., Okajima, K., Cheok, A. D., Petit, O., & Michel, C. (2016). Eating with our eyes: From visual hunger to digital satiation. *Brain and Cognition*, *110*. https://doi.org/10.1016/j.bandc.2015.08.006

Ursoniu, S., Serban, C. L., Giurgi-Oncu, C., Rivis, I. A., Bucur, A., Papava, I., & Bredicean, A. C. (2022). Validation of the Romanian Version of the Social Media Addiction Scale-Student Form (SMAS-SF) Among Undergraduate Medical Students. *Neuropsychiatric Disease and Treatment*, *18*. https://doi.org/10.2147/NDT.S368476

Van Strien, T., Frijters, J. E. R., Bergers, G. P. A., & Defares, P. B. (1986). The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *International Journal of Eating Disorders*, *5*(2). https://doi.org/10.1002/1098-108X(198602)5:2<295::AID-EAT2260050209>3.0.CO;2-Ta

World Health Organization. (2020, April 29). *Healthy Diet*. https://www.who.int/news-room/fact-sheets/detail/healthy-diet

Yong, C., Liu, H., Yang, Q., Luo, J., Ouyang, Y., Sun, M., Xi, Y., Xiang, C., & Lin, Q. (2021). The relationship between restrained eating, body image, and dietary intake among university students in china: A cross-sectional study. *Nutrients*, *13*(3). https://doi.org/10.3390/nu13030990

Zaiamri, M., & Abidin, Z. (2023). *Exploring Social Media Addiction among Malaysian University Students Ketagihan Media Sosial dalam kalangan Pelajar Universiti di Malaysia: Suatu Kajian Penerokaan*. https://www.researchgate.net/publication/376597751

ORIGINAL ARTICLE

LeptoSafe: Integrating Digital Health Tools in Early Detection and Prevention of Leptospirosis

Muhammad Amirul Afiq Rosli[1], Irfan Sahak[1], Norsaremah Salleh[1], Samer Zein[2], Roslaili Khairudin[3], Amir Faisal Mohd Khairi[4], Mohd Arifin Kaderi[5], and Norafiza Zainuddin[5,*]

[1]Department of Computer Science, Kulliyyah of ICT, International Islamic University Malaysia, 50728 Kuala Lumpur, Malaysia
[2]Department of Computer Science, Faculty of Engineering & Technology, Birzeit University, Palestine
[3]Communicable Diseases Control Unit, Pahang State Health Department, Jalan IM 4, Bandar Indera Mahkota, 25582 Kuantan, Pahang, Malaysia
[4]Maran District Health Office, Level 3, Maran Federal Building, 26500 Maran, Pahang, Malaysia
[5]Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia

Keywords:
Leptospirosis; Early detection; Prevention; Digital health; Mobile app

ABSTRACT

Background: Leptospirosis is a significant zoonotic disease endemic to tropical, flood-prone regions such as Pahang, Malaysia, where limited public awareness, delayed risk recognition, and overlapping symptoms with other febrile illnesses contribute to substantial underreporting and disease burden. Challenges in early detection and prevention necessitate innovative digital solutions to improve public health outcomes. **Methods**: *LeptoSafe*, a mobile health application prototype, was developed through a user-centred, iterative design approach involving collaboration with allied health experts and the Malaysian Ministry of Health. The app integrates evidence-based literature and real-time leptospirosis case data. Key components include a validated ten-item self-screening questionnaire for personalized risk assessment, GIS-enabled dynamic hotspot mapping with push notifications, educational infographics and videos, and a locator for nearby health facilities using Google Maps API. The platform supports bilingual access (English and Malay). **Results**: *LeptoSafe* integrates four core functionalities: (1) interactive self-assessment delivering individualized leptospirosis risk stratification; (2) real-time geospatial visualization of confirmed case clusters enhancing situational awareness; (3) comprehensive multimedia educational content tailored to local epidemiology; and (4) a health facility locator improving care accessibility. The prototype effectively addresses critical gaps in early symptom recognition, health knowledge dissemination, and timely healthcare engagement, particularly among vulnerable populations exposed to environmental and occupational risks. It represents a novel interdisciplinary advancement by integrating mobile technology, geographic information systems, and allied health expertise into a unified digital platform for zoonotic disease management. Moreover, *LeptoSafe* is among the first applications to combine real-time epidemiological data, risk self-assessment, and targeted education for leptospirosis prevention. Its modular and scalable design offers a replicable model for digital health interventions across neglected tropical diseases globally, enhancing surveillance capacity and public empowerment. **Conclusion**: *LeptoSafe* demonstrates the feasibility and impact potential of an integrated digital health platform to enhance leptospirosis control through empowerment, early detection, and prevention. Future directions include field deployment, rigorous evaluation of behavioural and knowledge outcomes, and integration with national surveillance systems to strengthen public health responses.

INTRODUCTION

Leptospirosis in Malaysia: Epidemiology, Risk Factors, and Challenges in Diagnosis and Control

Leptospirosis has emerged as a significant public health concern in Malaysia, recognized as an endemic zoonotic disease with a rising incidence and considerable impact (Ab Kadir et al., 2023; Ahmad Zamzuri et al., 2023; Lea et al., 2025; Mohd Yusof et al., 2025). In Malaysia, the disease burden has markedly escalated over the past two decades, with reported cases increasing from 1.45 to 25.94 per 100,000 population between 2005 and 2014 (Ab Kadir et al., 2023; Kamal and Foo, 2023). Despite leptospirosis being endemic in Malaysia, challenges persist in its diagnosis and management due to clinical overlap with other febrile illnesses and the limitations of diagnostic tools. A significant proportion of leptospirosis cases remain asymptomatic, particularly among urban sanitation workers, complicating surveillance and control efforts (Mohd Yusof et al., 2025). The disease is often misdiagnosed as dengue, malaria, or other febrile illnesses due to overlapping symptoms and limited diagnostic capacity in certain regions (Fornazari et al., 2021; Uribe-

* *Corresponding author.*
E-mail address: znorafiza@iium.edu.my

Restrepo et al., 2024, Esteves et al., 2025; Lea et al., 2025). This highlights the urgent need to better understand the dynamics of leptospirosis transmission, at-risk populations, and effective control measures to reduce morbidity and mortality.

Epidemiological data indicate notable spatial clustering in regions such as Sarawak, Kelantan, and Negeri Sembilan, with incidence rates typically peaking after flood events. This underscores the significant influence of hydroclimatic factors on disease transmission (Azman et al., 2025). Surveillance has shown that adult males are disproportionately affected by the infection, with mortality rates varying depending on the severity of the disease and the timeliness of diagnosis (Nazli et al., 2020). High-risk occupational groups, such as town service workers, oil palm plantation labourers, wet market vendors, and urban sanitation workers, exhibit elevated seroprevalence and infection risks due to their frequent exposure to contaminated environments and animals (Hanapi et al., 2021; Rahman et al., 2018; Samsudin et al., 2018; Jeffree et al., 2020; Mohd Yusof et al., 2025; Lea et al., 2025). Additionally, refugee and marginalized communities are particularly vulnerable due to poor living conditions (Hanapi et al., 2021; Azman et al., 2025).

Environmental factors, such as proximity to water bodies, flood exposure, inadequate sanitation, rapid urbanization, and rodent infestations, are consistently linked to heightened transmission risk (Pui et al., 2017; Mohd Taib et al, 2023; Suut et al., 2023; Baharom et al., 2024; Azman et al., 2025). Research has shown that specific land use types such as agricultural areas and low-lying regions along with urban features like mixed commercial-residential buildings, play a role in facilitating the circulation of the pathogen (Baharom et al., 2024; Azman et al., 2025; Lea et al., 2025). Rodents, particularly *Rattus* species, serve as primary reservoirs, harbouring pathogenic *Leptospira* spp. in both urban and rural settings (Mohd Taib et al., 2023; Suut et al., 2018; Laun et al., 2019). Recreational activities involving water contact, such as hiking and water sports, have also been linked to leptospirosis outbreaks (Lea et al., 2025). The detection of pathogenic *Leptospira* spp. in recreational waters and soils emphasizes the public health risks associated with these activities (Lea et al., 2025; Mohd Yusof et al., 2025).

Despite advances in diagnostic techniques, early detection of leptospirosis remains a challenge due to its nonspecific clinical manifestations and the limitations of traditional tests such as the microscopic agglutination test (MAT) (Jeffree et al., 2020) and enzyme-linked immunosorbent assay (ELISA) (Samsudin et al., 2018). Emerging diagnostic modalities, including molecular techniques like PCR (Atil et al., 2020) and machine learning models applied to microscopic agglutination images, show promise in improving diagnostic speed and accuracy (Rahmat et al., 2020). Predictive modelling that incorporates meteorological data and stochastic transmission models is also being employed to forecast outbreaks and identify potential hotspots (Rahmat et al., 2020; Rahim et al., 2023; Rahmat et al., 2021; Ramali et al., 2023).

Cross-Platform Mobile Tools for Leptospirosis Education

Research on cross-platform mobile applications for self-assessment and risk education has become a critical area of focus, driven by the growing reliance on mobile health technologies to improve disease prevention and management. Over the past decade, mobile health applications have evolved from basic educational tools to sophisticated platforms that integrate features such as interactive tutorials, symptom checkers, and location-based alerts (Kadir et al., 2025; Mohd Yusof et al., 2025; Suhaimi et al., 2025). This technological progression aligns with the increasing global burden of infectious diseases, including leptospirosis, highlighting the need for accessible and user-friendly digital interventions to enhance public health outcomes. The practical importance of these applications lies in their ability to provide timely risk calculators and customized educational content, empowering individuals to engage in self-directed learning and make informed decisions about their health.

Despite the widespread adoption of mobile health apps, several challenges persist in effectively integrating disease-specific features, such as symptom checkers and location-based alerts tailored to leptospirosis risk. This highlights a significant knowledge gap in the field. While existing literature frequently addresses general mobile health interventions or clinical decision support systems, comprehensive evaluations of cross-platform applications combining interactive educational content with real-time risk assessment tools for leptospirosis remain scarce.

Therefore, this research aims to develop *LeptoSafe*, a mobile application integrating self-screening, real-time hotspot mapping, and educational interventions to enhance early detection and prevention, focusing on its ability to deliver both educational and risk-assessment functionalities for leptospirosis.

METHODOLOGY

Study Design

This project employed a user-centred, iterative development framework to create *LeptoSafe*, a mobile application for leptospirosis early detection, prevention, and education. Design activities were conducted in close collaboration with Ministry of Health and allied health experts from the Kulliyyah of Allied Health Sciences

(KAHS), IIUM.

Application Development Approach

A Rapid Application Development (RAD) model guided the system's creation, emphasizing prototyping and incremental refinement under a six-month timeline (Pereira et al., 2022). Flutter is used for the frontend development of the prototype. It is a cross-platform mobile app framework that enables building iOS and Android apps from a single codebase. Backend services such as Firebase provide real-time database and authentication, while the Google Maps API is integrated to display leptospirosis hotspot locations. The development involves four phases described below:

Phase 1: Requirements Planning
–Open-ended interviews with allied health experts were conducted to elicit core functional requirements (self-screening, hotspot alerts, educational content, facility locator) and data sources.
– The project scope, success criteria, and technical constraints (data availability, geospatial accuracy, multimedia performance) were defined in this phase.

Phase 2: User Design
– This phase involves designing the low- and high-fidelity wireframes in Flutter, drawing on Material Design principles to ensure cross-platform consistency and accessibility.

Phase 3: Construction
– The frontend modules and backend services were implemented using Flutter in this phase. Firebase is also used for authentication, real-time database updates, and push notifications.
– We integrated QGIS for hotspot geoprocessing and the Google Maps API for facility lookup.
– Construction of content such as infographics and video resources, sourced from peer-reviewed literature and Ministry of Health reports was performed in this phase.

Phase 4: Cutover Preparation
– This phase involves finalizing the prototype, including automated data synchronization, user risk-scoring algorithms, and multilingual support (English, Malay).
– However, technical documentation, user guides, and test scripts for pilot deployment will be prepared in the next phase.

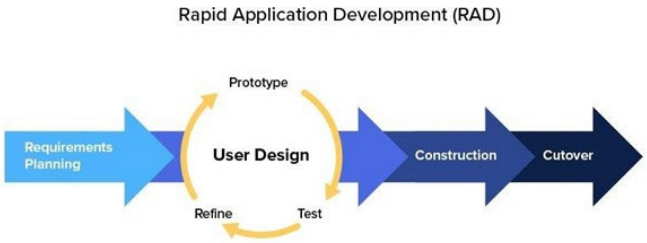


Figure 1: *Rapid Application Development Model*

Data Sources and Content Criteria

Leptospirosis Case Data
Aggregated from Malaysia's Ministry of Health for hotspot mapping. Cases older than one year were excluded to ensure timeliness.

Literature-Based Content
Educational materials (infographics, videos) were curated from PubMed, Web of Science, and Google Scholar; only content describing clinical symptoms, transmission routes, or preventive measures was included.

Self-Screening Survey Design

A ten-item questionnaire was designed to assess recent exposure history, symptom presence, and behavioural risk factors. Each response is scored and mapped to one of four risk levels (low, moderate, high, critical), triggering personalized recommendations ranging from self-monitoring advice to immediate clinical referral.

Hotspot Mapping Module

Geospatial Processing
QGIS workflows aggregate confirmed case coordinates into dynamic risk layers.

Real-Time Updates
A Firebase cloud function imports new case reports daily, refreshes map tiles, and issues location-based push notifications when users enter areas with elevated case density.

Logical Design

The use case diagram of *LeptoSafe* is designed as shown in Figure 2. User shall be able to login or register if there is no existing account, view hotspot location, do self-checking survey and view risk level along with personalized recommendation. Additionally, user can view leptospirosis symptom, treatment, and educational awareness content, find out nearby health facilities, and change language settings. The admin is responsible to manage hotspot location by updating the database.

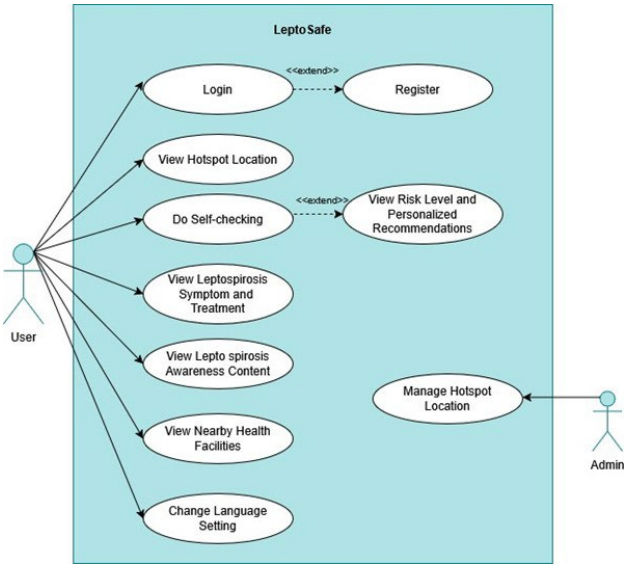


Figure 2: Use Case Diagram

Figure 3 shows the activity diagram of the user login and registration process. A user can either register by completing a form or log in by entering their credentials. The system then verifies the login details, if the information given is correct, the app dashboard will be displayed, or otherwise, the user is prompted to try again.

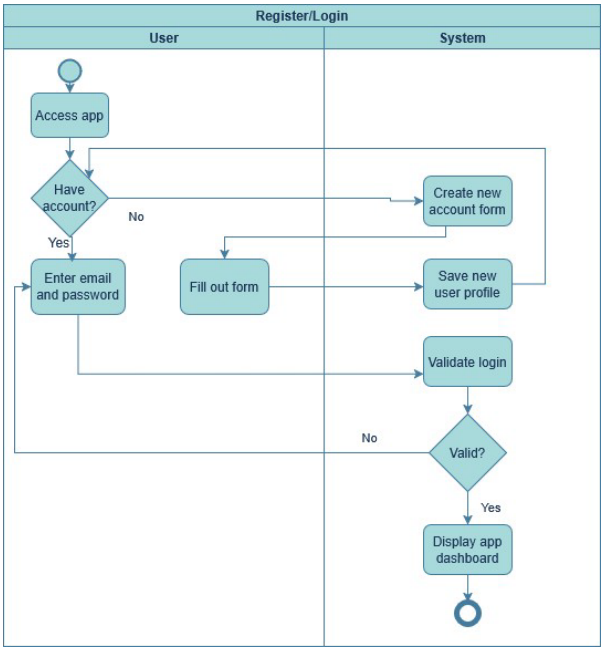


Figure 3: Register/Login Activity Diagram

The activity diagram in Figure 4 shows the activity when a user performing a self-check. The user navigates to the self-checking page, clicks "Self-checking," answers a questionnaire, and clicks "Submit." The system then displays the questionnaire, evaluates the user's responses, generates and saves a health record, and finally displays the health record to the user.

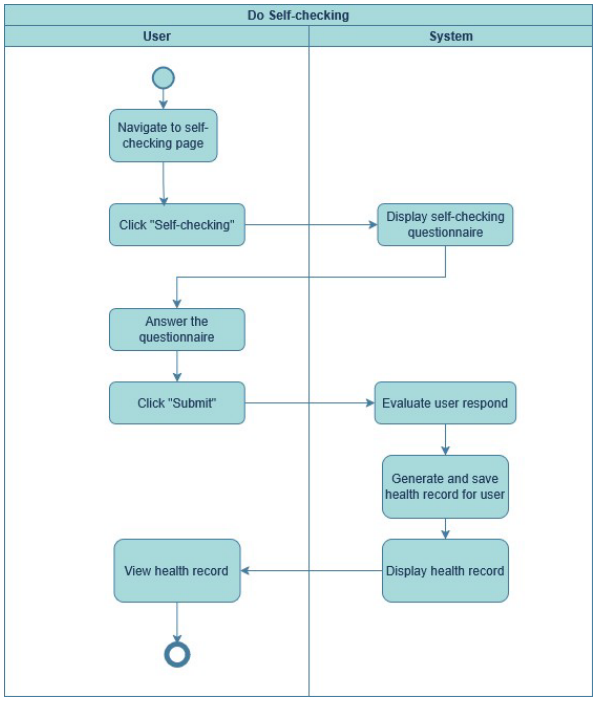


Figure 4: Do Self-checking Activity Diagram

The activity diagram shown in Figure 5 depicts the process of a user accessing their health records. When the user clicks "Health Record", the system retrieves and displays the available records. The user then selects a specific record, after which the system presents a personalized recommendation for the user to view.

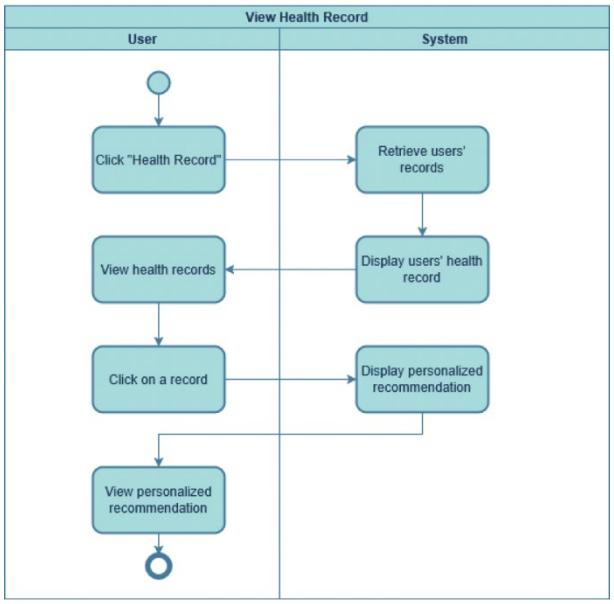


Figure 5: View Health Record Activity Diagram

The activity diagram in Figure 6 illustrates how a user accesses leptospirosis hotspot information. The user selects "Hotspot", prompting the system to display the hotspot page. After the user enters a desired location, the system presents nearby leptospirosis hotspots for viewing.

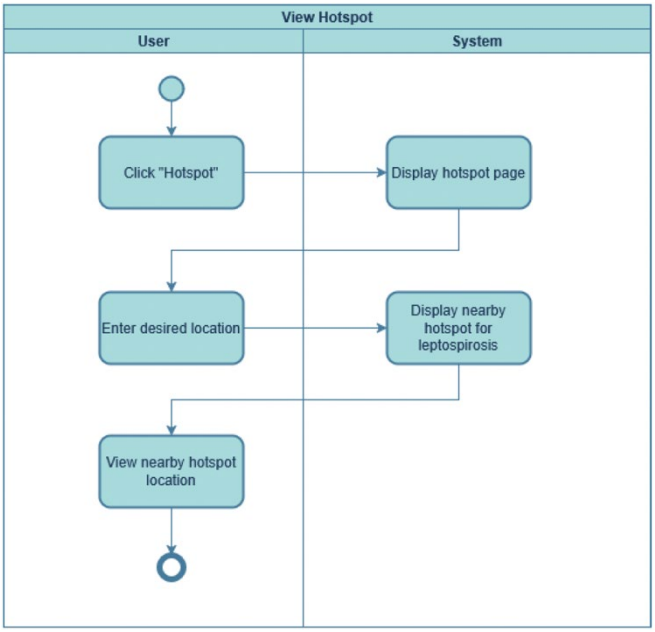


Figure 6:: View Hotspot Activity Diagram

The activity diagram in Figure 7 illustrates the process of viewing awareness content. When the user clicks "Awareness", the system displays the awareness page. The user then selects either the "Infographic" or "Video" tab, prompting the system to retrieve and display the selected content for the user to view.

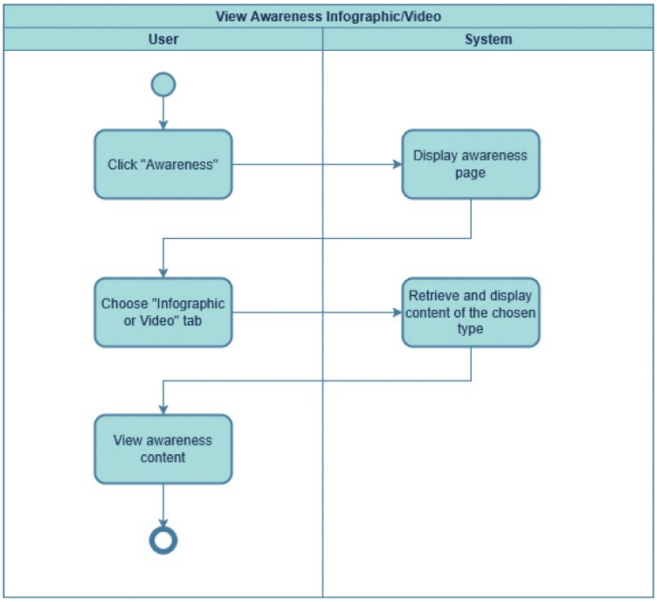


Figure 7: View Educational Awareness Content Activity Diagram

A sequence diagram is a type of interaction diagram that visually represents how objects in a system interact over time (Booch et al., 2005). It shows the order of messages exchanged between objects, typically illustrating the flow of a specific scenario or use case. The sequence diagram in Figure 8 shows the self-checking process. A user starts self-checking via the app, answers a questionnaire, then the app evaluates responses, generates then saves a health record in the database, and displays it. The user can then view health records and click one to see risk level and personalized recommendations.

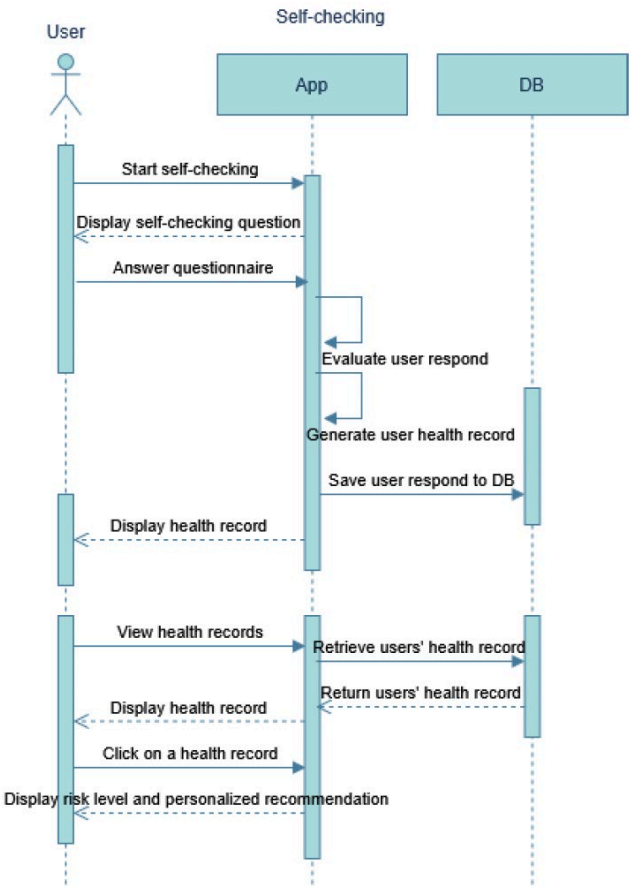


Figure 8: Self Checking Sequence Diagram

RESULTS

In this Section, a high-fidelity prototype of the mobile application (*LeptoSafe*) is presented. Figure 9 shows the landing page that will be displayed when *LeptoSafe* prototype is executed. User is then required to sign in using their registered email.

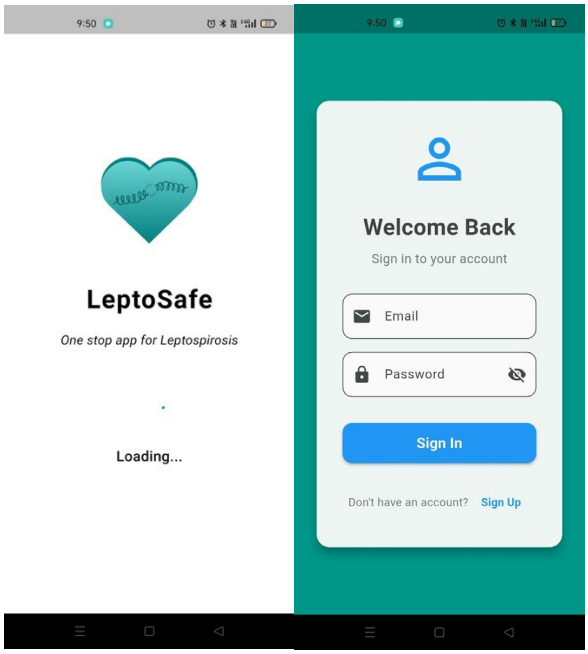


Figure 9: Landing and welcome page

Once the user has logged in, the dashboard page shown in Figure 10 will be displayed. It shows the main features that user can benefit from. For example, when clicking on the "Health Record" icon, their health information will be shown.

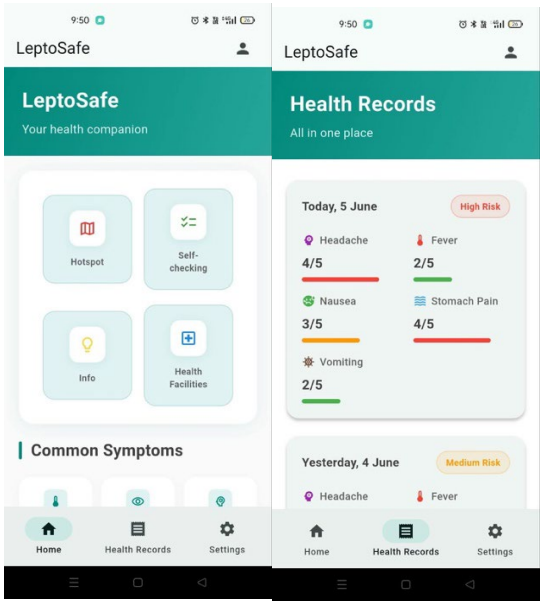


Figure 10: Dashboard and health record page

The self-screening can be done by user to rate the severity level of Leptospirosis symptoms. User will be prompted with a total of ten questions using 5-point Likert Scale (see Figure 11). From the dashboard, user is also allowed to update their profile page.

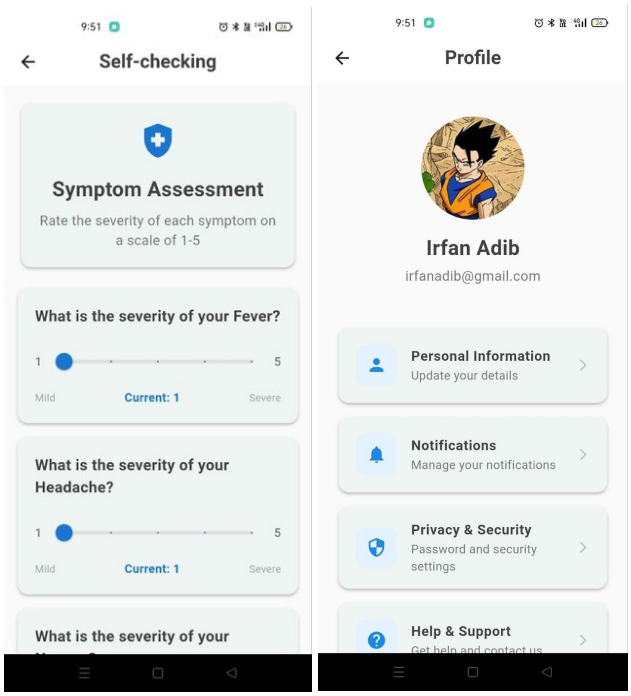


Figure 11: Self-screening and Profile page

Another main feature of the prototype is to show the hotspot location for the user to check the location of leptospirosis cases. They can also view the available health facilities as can be seen in Figure 12.

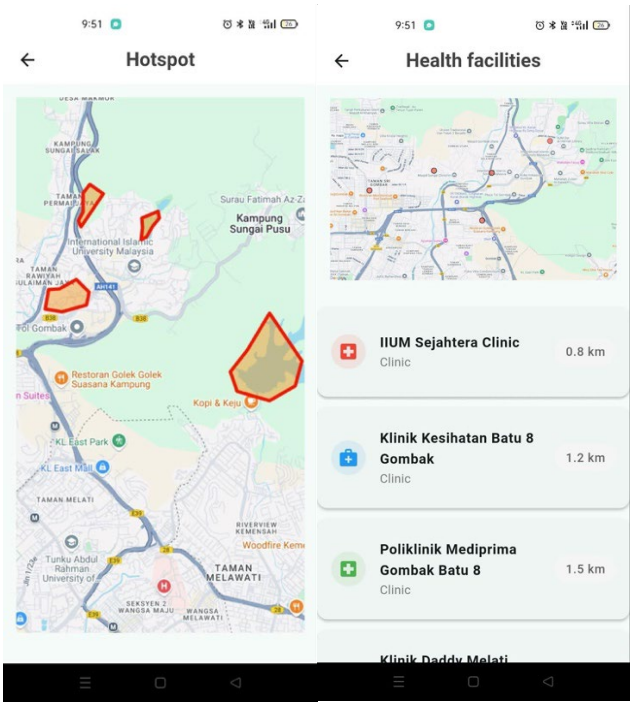


Figure 12: Hotspot map and Health Facilities

System Implementation

LeptoSafe's core modules were implemented in a cross-platform Flutter application with Firebase and QGIS integration. Table 1 summarizes the features. For the prototype development, we focused on development of three modules: i) User Authentication; ii) Self-Screening Survey, and iii) Educational Content.

Table 1: Modules and Descriptions

Module	Description
User Authentication	Registration and login with email/password, Firebase Authentication integration
Self-Screening Survey	Ten-item questionnaire with algorithmic risk scoring and personalized recommendation output
Hotspot Mapping	Real-time GIS visualization of confirmed case clusters via QGIS; push notifications for alerts
Educational Content	Infographics and short videos on symptoms, transmission, and prevention
Nearby Health Facility Locator	Google Maps API integration to display clinics and hospitals within 5 km of user location
Multilingual Support	English and Malay language options for all user interface elements

DISCUSSION

LeptoSafe represents a promising advancement in

integrating digital health tools for leptospirosis early detection, prevention, and education. Its core features—self-screening survey, real-time hotspot mapping, and educational content—address key challenges in leptospirosis management, especially in resource-limited and flood-prone areas where healthcare access may be limited (Mohd Yusof et al., 2025; Kamal & Foo, 2023). The personalized risk assessment facilitated by the self-screening survey empowers users to make informed healthcare decisions, which is vital in settings where delayed diagnosis often worsens outcomes (Fornazari et al., 2021; Jeffree et al., 2020).

The dynamic hotspot mapping with push notifications aligns with recent advances in using geospatial data to predict and monitor leptospirosis incidence, thereby improving situational awareness (Ab Kadir et al., 2023; Ramali et al., 2023). Such spatial epidemiology tools have been shown to enhance public health responses by identifying high-risk areas and facilitating targeted interventions (Rahim et al., 2023). The integration of a health facility locator further supports timely access to care, complementing existing gaps in healthcare infrastructure within endemic regions (Ahmad Zamzuri et al., 2023).

Educational modules utilizing multimedia formats engage diverse users effectively, closing knowledge gaps that contribute to the underreporting and misdiagnosis of leptospirosis (Lea et al., 2025; Baharom et al., 2024). This is particularly crucial in tropical settings with overlapping febrile illnesses such as dengue and malaria, where confusion in symptom recognition complicates clinical management (Esteves et al., 2025; Uribe-Restrepo et al., 2024).

Importantly, *LeptoSafe* exemplifies the benefits of interdisciplinary collaboration between computer scientists, allied health experts, and public health authorities, a model advocated in other digital health interventions for zoonotic and neglected tropical diseases (Suhaimi et al., 2025; Mohd Yusof et al., 2025). Its modular, scalable architecture offers potential adaptability to other infectious diseases, enhancing its value beyond leptospirosis.

While the prototype demonstrates technological feasibility and aligns with epidemiological evidence, field validation is essential. Future work should rigorously assess user engagement, knowledge gains, behaviour changes, and epidemiological impacts in diverse populations. Furthermore, integration with national surveillance and reporting systems could transform *LeptoSafe* from a standalone tool to a critical component of Malaysia's leptospirosis control strategy (Kadir et al., 2025; Kamal & Foo, 2023).

CONCLUSION

LeptoSafe presents a feasible and innovative solution for the early detection, prevention, and education of leptospirosis. The app's integrated features—self-screening, real-time hotspot mapping, and educational content—demonstrate its potential to address key public health challenges associated with the disease. While the initial prototype has shown positive usability and user engagement, further research is needed to evaluate its impact on health behaviour and its integration into national surveillance systems. The continued development and testing of *LeptoSafe*, informed by field data, will be crucial in ensuring its effectiveness and scalability in the fight against leptospirosis.

ACKNOWLEDGEMENT

This research was not funded by any grant.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author(s) used Perplexity AI to draft the manuscript text and format the citations. The author(s) subsequently reviewed, revised, and approved all content, and accept full responsibility for the final manuscript.

REFERENCES

Ab Kadir, M. A., Abdul Manaf, R., Mokhtar, S. A., & Ismail, L. I. (2023). Spatio-temporal analysis of leptospirosis hotspot areas and its association with hydroclimatic factors in Selangor, Malaysia: Protocol for an ecological cross-sectional study. *JMIR Research Protocols*, 12, e43712. https://doi.org/10.2196/43712

Ahmad Zamzuri, M. 'A. I., Mohd Shah, S., Junadi, N., Abd Latif, M. A., Rosly, K. A., Abd Majid, F. N., Dapari, R., Mihat, M., Ibrahim, S. S., Ismail, M., Abd Aziz, S., Hassan, M. R., Rejali, L., Abdullah, Z., & Mohd Isa, A. M. (2023). Demystifying re-emergence zoonotic threat: A 5-year epidemiology of leptospirosis in Negeri Sembilan, Malaysia. *Clinical Epidemiology and Global Health*, 24, 101432. https://doi.org/10.1016/j.cegh.2023.101432

Atil, A., Jeffree, M. S., Syed Abdul Rahim, S. S., Hassan, M. R., Lukman, K. A., & Ahmed, K. (2020). Occupational determinants of leptospirosis among urban service workers. *International Journal of Environmental Research and Public Health*, 17(2), 427. https://doi.org/10.3390/ijerph17020427

Azman, N. A., Rasudin, N. S., Atan, E. H., Ghafar, N. A., Ismail, R., & Chandra, N. D. S. (2025). The association of climatic pattern and leptospirosis cases in Malaysia from 2013 to 2021. *Malaysian Journal of Pathology*, 47(1). https://www.mjpath.org.my/2025/v47n1/climatic-pattern-and-leptospirosis.pdf

Baharom, M., Ahmad, N., Hod, R., Ja'afar, M. H., Arsad, F. S., Tanggang, F., Ismail, R., Mohamed, N., Mohd Radi, M. F., & Osman, Y. (2024). Environmental and occupational factors associated with leptospirosis: A systematic review. *Heliyon*, 10, e23473. https://doi.org/10.1016/j.heliyon.2023.e23473

Booch, G., Rumbaugh, J., & Jacobson, I. (2005). *The Unified Modeling Language User Guide* (2nd ed.). Addison-Wesley.

Esteves, S. B., de Oliveira, L. M., Guilloux, A. G. A., Cortez, A., de Masi, E., Ferreira, I. M. R., de Lima, E. M., Ramalho, G. Y., Luna, P. C., Ferreira Neto, J. S., Heinemann, M. B., & Miotto, B. A. (2025). Into the spotlight: A spatial study of potentially underreported leptospirosis among dengue-negative patients in São Paulo city, Brazil. *PLoS Neglected Tropical Diseases*, 19(3), e0012888. https://doi.org/10.1371/journal.pntd.0012888

Fornazari, F., Richini-Pereira, V. B., Joaquim, S. M., Nachtigall, P. G., & Langoni, H. (2021). Leptospirosis diagnosis among patients suspected of dengue fever in Brazil. *Journal of Venomous Animals and Toxins including Tropical Diseases*, 27, e20200118. https://doi.org/10.1590/1678-9199-JVATITD-2020-0118

Hanapi, I. R. M., Sahimin, N., Maackara, M. J. B., Annisa, A. S., Mutalib, R. N. S. A., Lewis, J. W., Behnke, J. M., Lau, Y. L., & Zain, S. N. M. (2021). Prevalence of anti-Leptospira antibodies and associated risk factors in the Malaysian refugee communities. *BMC Infectious Diseases*, 21, 1128. https://doi.org/10.1186/s12879-021-06830-0

Hii, K. C., Robie, E. R., Saihidi, I., Berita, A., Alarja, N. A., Xiu, L., Merchant, J. A., Binder, R. A., Goh, J. K. T., Guernier-Cambert, V., et al. (2021). Leptospirosis infections among hospital patients, Sarawak, Malaysia. *Tropical Diseases, Travel Medicine and Vaccines*, 7, 32. https://doi.org/10.1186/s40794-021-00154-2

Jeffree, M. S., Mori, D., Yusof, N. A., Atil, A. B., Lukman, K. A., Othman, R., Hassan, M. R., Suut, L., & Ahmed, K. (2020). High incidence of asymptomatic leptospirosis among urban sanitation workers from Kota Kinabalu, Sabah, Malaysian Borneo. *Scientific Reports*, 10, 19442. https://doi.org/10.1038/s41598-020-76595-0

Kamal, N., & Foo, D. C. W. (2023). Spatio-temporal analysis of leptospirosis hotspot areas and its association with hydroclimatic factors in Selangor, Malaysia. *JMIR Research Protocols*, 12(1), e43712. https://doi.org/10.2196/43712

Kadir, M. A. A., Manaf, R. A., Mokhtar, S. A., & Ismail, L. I. (2025). Identifying leptospirosis hotspots in Selangor: Uncovering climatic connections using remote sensing and developing a predictive model. *International Journal of Environmental Research and Public Health*, 20(5), 3651.

Lau, C. L., Smythe, L. D., & Craig, S. B. (2019). Association of rodent-borne Leptospira spp. with urbanization and environmental factors. *PLoS Neglected Tropical Diseases*, 13(2), e0007141. https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0007141

Laun, P., et al. (2019). Association of rodent-borne Leptospira spp. with urbanization and environmental factors. *PLoS Neglected Tropical Diseases*, 13(2), e0007141. https://doi.org/10.1371/journal.pntd.0007141

Lea, J. S. X., Reduan, M. F. H., Choong, S. S., Kamaruzaman, I. N. A., Ooi, P. T., AbuBakar, S., Loong, S. K., & Rahman, M. S. A. (2025). Leptospirosis in humans and animals in Malaysia: A review from 1976 to 2023. *Veterinary World*, 18(3), 673–685. https://doi.org/10.14202/vetworld.2025.673-685

Mohd Taib, F. S., et al. (2023). Leptospirosis: An insight into community structure of small mammal's reservoir hosts in Malaysia. *Tropical Biomedicine*, 37(1), 142-154. https://msptm.org/files/Vol37No1/142-154-Mohd-Taib-FS.pdf

Mohd Yusof, M. S., Ahmed, N., Bakhir, S. N. I., Shanghari, R. V. S., & Fauzi, M. H. (2025). The prevalence of leptospirosis infections among humans in Malaysia: A systematic review and meta-analysis. *Tropical Biomedicine*, 42(2), 201-212. https://doi.org/10.47665/tb.42.2.013

Nazli, A., Rosmiati, A., & Shahril, M. R. (2020). Clinical presentations and predictors of mortality for leptospirosis in Larut, Matang, and Selama districts, Perak, Malaysia. *Medical Journal of Malaysia*, 75(1), 56-64. Retrieved from https://www.e-mjm.org/2020/v75n1/leptospirosis.pdf

Pereira, R., Moreira, P., & Hagino, T. (2022). *Rapid application development with OutSystems: Create applications with OutSystems up to seven times faster than with traditional technologies*. Packt Publishing.

Pui, C. F., Bilung, L. M., Apun, K., & Su'ut, L. (2017). Diversity of Leptospira spp. in rats and environment from urban areas of Sarawak, Malaysia. *Journal of Tropical Medicine*, 2017, 3760674. https://doi.org/10.1155/2017/3760674

Rahim, S. M., Zulkarnain, N., & Jamal, S. (2023). Leptospirosis modelling using hydrometeorological indices and random forest classifier in Kelantan, Malaysia. *International Journal of Environmental Research and Public Health*, 20(6), 1234. https://pubmed.ncbi.nlm.nih.gov/36719482/

Rahmat, F., Zulkafli, Z., Ishak, A. J., Noor, S. B. M., & Yahaya, H. (2021). Prediction model of leptospirosis occurrence for Seremban (Malaysia) using meteorological data. *International Journal of Integrated Engineering*, 13(1), 60-69. https://climahealth.info/resource-library/prediction-model-of-leptospirosis-occurrence-for-seremban-malaysia-using-meteorological-data/

Rahmat, F., Zulkafli, Z., Ishak, A. J., Noor, S. B. M., Yahaya, H., & Masrani, A. (2020). Exploratory data analysis and artificial neural network for prediction of leptospirosis occurrence in Seremban, Malaysia based on meteorological data. *Frontiers in Earth Science*, 8, 377. https://doi.org/10.3389/feart.2020.00377

Ramali, N., Ibrahim, A., Syafiq, F. M., & Norazian, M. (2023). Spatio-temporal analysis of leptospirosis hotspot areas and its association with hydroclimatic factors in Selangor, Malaysia. *International Journal of Public Health Research*, 12(1).

Suhaimi, N., Ahmad, N., Zulkifli, R., & Hassan, R. (2025). Infectious disease traveler mobile application based on need assessment analysis in Malaysian students. *Makara Journal of Health Research*, 29(1).

Suut, L., Mazlan, M. N. A., Arif, M. T., Katip, T., Aliza, A. R. N., & Haironi, Y. (2018). Serovar diversity of Leptospira sp. infecting wild rodents in Sarawak, Malaysia. *Tropical Biomedicine*, 35(1), 252-258. https://msptm.org/files/Vol35No1/252-258-Suut-L.pdf

Uribe-Restrepo, P., Perez-Garcia, J., Arboleda, M., Munoz-Zanzi, C., & Agudelo-Florez, P. (2024). Clinical presentation of human leptospirosis in febrile patients: Urabá, Colombia. *PLoS Neglected Tropical Diseases*, 18(9), e0012449. https://doi.org/10.1371/journal.pntd.0012449

REVIEW ARTICLE

Allergy Prevention Mechanisms of Human Milk Proteomes in Infants: A Scoping Review

Radiah Abdul Ghani[1,*], Nurul Faqihah Shaharul[1], Eka Norhayati[2]

[1]Department of Biomedical Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2]Department of Midwifery, Faculty of Health Science, Universitas Alma Ata, Yogyakarta, Indonesia

Keywords:
human milk; proteomes; allergy; infants

ABSTRACT

Background: There has been an evident rise in global prevalence of allergic diseases, such as food allergies, eczema, asthma, and allergic rhinitis but most notably in infants in recent times. Human milk contains diverse bioactive compounds that modulate immunity and protect against allergies. Of the main bioactive components of human milk, the human milk proteome, a broad spectrum of proteins as well as peptides has emerged as a key agent in modifying allergic reactions.This study aims to explore the relationship between the human milk proteome and the development of allergies in infants, focusing on their mechanisms in preventing allergen sensitivity **Methods:** A comprehensive search of peer-reviewed articles published between 2015-2025 was conducted in PubMed, Scopus, and ScienceDirect using the keywords: ("human milk" OR "breast milk") AND ("proteome" OR "proteins") AND ("allergy" OR "hypersensitivity") AND ("infant" OR "newborn"). Of 533 articles retrieved, 26 passed screening, and 4 met the inclusion criteria, as illustrated in the PRISMA-ScR flow diagram. **Results:** It was found that there are seven proteomes of human milk that contain immunological components which are IgG, secretory IgA (SIgA), Vitronectin, CD81, C4A, IgG4 and interleukin-10 (IL-10). These components are involved in different mechanisms that prevent sensitivity to allergens such as transferring passive immunity that blocks allergen sensitization and promote immune tolerance by modulating inflammatory responses. The types of allergies included in this study are food allergies, atopic dermatitis, allergic rhinitis, asthma, allergic sensitization, eczema and respiratory allergies such as pollen allergies. **Conclusion:** The study highlights the imperative role of human milk proteomes such as immunoglobulins, sIgA, and IL-10 in lowering the risk of developing allergic illnesses like asthma, eczema, and allergic rhinitis by inhibiting allergic sensitization, preventing infections, and triggering a synergistic immune response. Further studies will be needed to investigate the synergistic roles of these bioactive components and their sustained effects on allergy prevention to improve global public health.

INTRODUCTION

The increasing prevalence of allergies, including food allergies, eczema, allergic rhinitis, and asthma, highlights the need for a deeper understanding of early-life factors influencing immune development. In accordance with the 2011 White Book on Allergy, the incidence of allergic disorders is experiencing a significant increase on a global scale, encompassing both industrialized and emerging nations, with the prevalence rates of asthma fluctuating between 1% and 20%, allergic rhinitis ranging from 1% to 18%, and dermatological allergies varying from 2% to 10% across diverse demographic groups (Gutowska-Ślesik *et al.*, 2023). More recent global reports from the Global Allergy and Asthma European Network (GA²LEN), the World Allergy Organization (WAO), and analyses from the Global Burden of Disease (GBD 2021) confirm that allergic diseases continue to pose a major public health challenge. The GBD 2021 estimates indicated that approximately 260 million people were affected by asthma and 129 million by atopic dermatitis worldwide in 2021, with projections showing further increases by 2050 (Oh *et al.*, 2025). The growing prevalence of allergies is particularly noticeable in children, who have been affected most by this trend for the last two decades. Allergy prevention targets the prevention of the development of allergic diseases and the reduction of the risk of its occurrence. Allergy prevention entails measures that are directed at reducing exposure to allergens and enhancing the immune system's capacity to distinguish between dangerous and safe agents. Allergy prevention is particularly influenced by determinants that occur in early life, including breastfeeding with a

[*]*Corresponding author.*
E-mail address: radiah@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

modulating effect on the immune system, which lowers the risk of allergic sensitization (Ballard & Morrow, 2013; ASCIA, 2020). Human milk is widely acknowledged for its protective role against allergies, attributed to its rich composition of bioactive components (Ballard & Morrow, 2013).

Among these, it is the human milk proteome which is the most dynamic collection of proteins and peptides that is emphasized due to its immunomodulatory characteristics. These proteomes assist in the maturation of the infant's immune response, foster tolerance towards allergens and modify gut microbiomes. For example, Kull *et al.* (2010) showed that the risk of allergic sensitization and asthma in children who were exclusively breastfed for at least four months was lower than in those who were breastfed for less than four months.

The human milk proteome includes nutritional proteins such as caseins and whey proteins, immunological proteins like lactoferrin, lysozyme, and secretory immunoglobulin A (SIgA), as well as bioactive peptides that influence immune regulation. According to Zhu and Dingess (2019), this proteomic diversity allows human milk to adapt to the changing needs of the infant throughout lactation, providing essential support for growth and development. These components are thought to play a significant role in fostering immune tolerance, a process by which the infant's immune system learns to distinguish between harmful and harmless antigens.

Despite evidence suggesting a protective effect of breastfeeding on allergy risk, inconsistencies in study findings and limited understanding of the role of specific human milk proteomes present a significant knowledge gap. Clarifying these mechanisms is crucial to developing targeted nutritional strategies that could mitigate allergy risk in infants. This study aims to explore the relationship between the human milk proteome and the prevention of allergies in infants, with a focus on identifying specific proteomes and their mechanisms of action in preventing allergen sensitivity.

METHODOLOGY

This study employed a scoping review methodology to investigate the role of the human milk proteome in the prevention of allergies in infants. Scoping reviews were suitable for this purpose as they enabled the comprehensive mapping of diverse literature, methodologies and findings across broad research areas. This approach was guided by the framework of Arksey and O'Malley (2005) and adhered to the PRISMA-ScR checklist to ensure clear and thorough reporting (Tricco et al., 2018). This design allowed for a detailed analysis of existing evidence, identifying mechanisms, gaps, and potential clinical applications.

Stage 1: Identify the Research Questions

The research questions identified which are 1) Which proteomes in human milk contribute to prevention of allergen sensitivity in infants? and 2) What are the mechanisms of action of the proteomes in human milk in preventing allergen sensitivity in infants? These research questions were used to guide the search strategy and were further evaluated using the eligibility criteria to choose the most relevant articles for inclusion in this scoping review.

Stage 2: Identify Relevant Studies

Research articles were collected from a number of prominent databases, including PubMed, Scopus, and ScienceDirect. The search for articles was limited to those published in the past ten years, i.e., 2015 to 2025, to guarantee inclusion of the most current and pertinent data on the subject. Search keywords were drawn from subject-related keywords of the human milk proteome and its role in immune tolerance and allergy prevention in infants. Boolean operators such as "AND," "OR," and "NOT" were employed to combine and exclude terms to restrict the search for better retrieval and capture of relevant articles. These operators were applied to restrict the search to articles that met the inclusion criteria of human milk proteome and infant allergies. For example, the following phrases were used: ((("human milk" OR "breast milk") AND ("proteome" OR "proteins") AND ("allergy" OR "hypersensitivity") AND ("infant" OR "newborn"))). By utilizing these operators, the search was optimized to identify the most relevant studies while avoiding non-relevant articles.

Stage 3: Study Selection

The following studies were included if they met the criteria: (1) fulfill both objectives, (2) publications from the last 10 years, (3) articles written in English, and (4) full-text articles. Studies were excluded if they were unpublished hand-searched, grey literature, review articles, book chapters, conference abstracts, letters or if they fulfilled only one objective.

Stage 4: Charting the Data

The data were charted beginning with duplicate citations removed for independent title, abstract, and full-text

review based on the inclusion criteria. Titles and abstracts were screened independently and in parallel for potential inclusion. Potentially eligible sources were retrieved in full text. Full-text sources were screened independently against the inclusion criteria. The results of the search and inclusion of the study were reported in full in the final scoping review and summarized in a PRISMA-ScR flow diagram as shown below in **Figure 1**.

Stage 5: Collating, Summarizing, and Reporting Result

This stage describes the key steps of the review process. First, the study characteristics were summarized, including the author, year of publication, study title, and journal source. Next, the methods used in each study were outlined, covering the study design, research objectives, and approaches applied. Finally, the findings were synthesized according to the objectives of this review, highlighting the human milk proteomes identified in the selected studies, their mechanisms of action in preventing infant allergies, and the types of allergies discussed. An informal quality appraisal was conducted to provide a clearer understanding of the robustness and validity of the included studies. It involved assessing the clarity of the study's purpose, the appropriateness of its design, and potential sources of bias. Although no standardized tool or ranking system was used, this process allowed for contextualization of the findings and the identification of research gaps, which may inform future high-quality studies. Data extraction through this process was helpful in addressing the objectives and research questions of the scoping review. These tables provided a descriptive summary of the findings and collectively created a more comprehensive picture of the study.

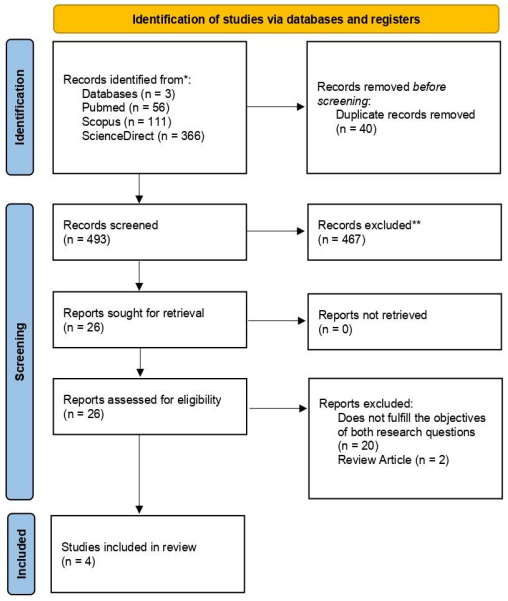


Figure 1: PRISMA ScR flow diagram

RESULTS

Study Characteristics

The selected articles were experimental research studies. This review included findings from several journals: Lupinek *et al.* (2019) in The Journal of Allergy and Clinical Immunology, Donald *et al.* (2024) in Cell Reports, Jiang *et al.* (2021) in Molecular Nutrition & Food Research, and Manti *et al.* (2016) in Immunobiology. All of the studies used human breast milk, although one employed a mouse model. Despite using different approaches, all studies confirmed the beneficial role of the human milk proteome in allergy prevention.

Human Milk Proteome Components and Their Mechanisms in Preventing Allergy

The included studies focused on specific components of human milk that contribute to allergy prevention. Donald *et al.* (2024) demonstrated that secretory IgA (sIgA) regulates the gut microbiome and prevents asthma. Jiang *et al.* (2021) identified multi-omics components such as vitronectin and CD81 that support immune development and reduce allergy risks. Lupinek *et al.* (2019) found that allergen-specific IgG in breast milk protects against respiratory and food allergies, while Manti *et al.* (2016) reported that milk-derived interleukin-10 (IL-10) promotes immune tolerance and alleviates cow's-milk-protein allergy. **Table 1** summarizes the mechanisms of action of these key proteomes in preventing specific types of allergies. This table is important for illustrating how each proteome component functions in relation to different allergic conditions. **Table 2** presents a consolidated summary of the main proteomes identified and their mechanisms of action. It provides a concise comparison that highlights how multiple human-milk proteins contribute synergistically to immune tolerance and reduced allergy risk in infants.

DISCUSSION

This study aimed to explore the role of proteomes present in human milk in preventing allergen sensitivity in infants. Specifically, the study was focused on identifying the specific proteomes of human milk involved in infant allergy prevention and identifying how these proteomes function to prevent allergen sensitivity. From the findings, human milk contains a wide range of bioactive molecules like immunoglobulins, SIgA, and IL-10, which are all crucial for ensuring the integrity of the immune system and causing resistance to allergies. This present scoping review presents important findings on the role of the human milk proteome in reducing inflammation, and gut health,

Table 1: Mechanisms of Proteomes in Preventing Allergies

Author & Year	Mechanism of Action	Type of Allergy
Donald *et al.*, 2024	Breast milk-derived secretory IgA (sIgA) modulates the neonatal gut microbiota by blocking immunogenic bacteria colonization (segmented filamentous bacteria, SFB), preventing overactive Th17 immune responses and reducing asthma susceptibility.	Asthma
Jiang *et al.*, 2021	Human milk proteome as factor 2 influence infant growth and susceptibility to allergy through interactions with infant metabolism and immune development. The multi-omics components co-regulate infant health outcomes.	Eczema, Asthma, Rhinitis, and Atopic Dermatitis
Lupinek *et al.*, 2019	Mother-specific allergen-specific IgG antibodies are transmitted through placenta and breast milk to infant and transfer passive immunity that blocks allergen sensitization and reduces child IgEmediated allergy induction.	Respiratory Allergies like pollens
Manti *et al.*, 2016	Breastfeeding provides immunologically active components like IL-10, which promotes immune tolerance by modulating inflammatory responses (Th1/Th2 balance), resulting in reduced severity of allergic symptoms and improved disease outcome.	Cow's Milk Protein Allergy (CMPA)

Table 2: Summary of Key Findings

Proteomes	Mechanism of Action
Secretory IgA (sIgA)	Blocks gut colonization by immunogenic bacteria; suppresses Th17 responses
Vitronectin, CD81, C4A, IgG4	• Interact with metabolism and immune signals via multi-omics pathways • Influence growth and allergy susceptibility
IgG	• Transfers maternal allergen-specific IgG; prevents IgE sensitization • Passive immunity, prevents allergy onset
IL-10	• Balances Th1/Th2 responses; suppresses inflammation • Promotes immune tolerance

particularly in preventing allergies in infants. This review synthesizes such findings, identifies their mechanisms of action, and affirms the need for increased research to enhance our understanding of the interaction between human milk proteomes and the immune system.

Mechanisms of Proteomes in Preventing Allergies

The proteomes that can prevent allergies are proven by collecting data from research articles. Firstly, in a study by Donald *et al.* (2024) stated that SIgA organizes infant gut microbial populations by adhering to specific bacteria. It limits colonization by certain immunogenic bacteria such as Segmented Filamentous Bacteria (SFB), which if uncontrolled, could overstimulate the immune system. By controlling bacterial colonization, SIgA prevents premature or excessive immune stimulation. Without SIgA, bacteria such as SFB overgrow and are in close contact with the gut lining. This bloom results in early and increased activation of Th17 cells, which are immune cells involved in inflammation. Increased Th17 responses are involved in allergic inflammation and asthma. SIgA controls excessive expansion of pro-inflammatory immune-driving bacteria during early life. This balances the types of immune cells, including regulatory T cells (Tregs) that promote allergen tolerance. This is supported by the findings of Rahman *et al.* (2023), which emphasized that SIgA is beneficial for mucosal immunity and regulates

innate and adaptive immunity and, by extension, dampens allergic inflammation risk. These findings also indicate immunological importance of SIgA for stabilizing milieus of guts and preventing immune hyperresponsiveness. Selective arranging within gut microbes through SIgA follows intricate molecular mechanisms revealed within the KEGG pathway hsa04672 (Intestinal Immune Network for IgA Production), where SIgA is secreted in an active form from plasma cells within mucosal tissues, then adheres with help through the polymeric immunoglobulin receptor (pIgR). In a selective manner, the attachment denotes SIgA stringently regulates microbial populations, where good species are free towards dominating with potentially pro-inflammatory ones, for example, SFB, kept in line. In the absence of these, as in the case for the Th17 cell differentiation pathway (hsa04659), SFB overgrowth can potentially initiate IL-6, TGF-β mediated Th17 polarization, heading towards inflammatory cytokine production like IL-17A, IL-22. Chronic allergic inflammation results from pro-inflammatory cytokines, with a break in immune tolerance. Also, SIgA interaction with Tregs, as in the case for the TGF-beta signaling pathway (hsa04350), initiates IL-10, TGF-β secretion, heading towards a state of anergy/immunological suppression, Th1/Th2/Th17 balance. In short, SIgA, besides downregulating dysbiosis-derived inflammation, also imposes an early tolerant, homeostatic immunologic environment, with a decrease in risk for allergy through multiple complementary immunoregulatory mechanisms.

Moreover, in the study by Jiang *et al.* (2021), key proteomes such as Vitronectin, CD81, Complement C4A, Immunoglobulin γ4 (IgG4) are categorized as factor 2. Vitronectin is involved in cell adhesion and immune regulation. CD81 involved immune cell signaling and regulation. Complement C4A is a part of the complement system, and it supports immune defense and pathogen removal. Immunoglobulin γ4 an antibody subunit important to immune response. Factor 2 is inversely associated with αs1-casein, a well-known cow milk allergenic protein. Samples of Factor 2 milk have reduced levels of the major milk allergen αs1-casein. Lower exposure to the allergen, lowers the risk of sensitization and allergy in infants like eczema, asthma, rhinitis, and atopic dermatitis. The combination of immuneregulatory proteins supports the development of immune tolerance, meaning the infant's immune system learns to recognize harmless antigens without overreacting. Statistically, higher levels of Factor 2 proteome are linked with a 35% lower risk for infant allergy. These are supplemented with KEGG pathway integrations. In Complement and Coagulation Cascades (hsa04610), Vitronectin action stabilizes interaction with complement components with a modulatory influence on inflammatory activity. C4A also takes part in immune tolerance through processing and clearance of immune complexes. In Cell Adhesion Molecules (CAMs) (hsa04514), CD81 participates in tetraspan-enriched microdomains, modulating immune cell adhesion and immune cell signaling. CD81 and Vitronectin are also common in Antigen Processing and Presentation (hsa04612), with regulated antigen recognition and immune regulation being supplied. IgG4, although cataloged for no given KEGG pathway, has shown an action as a "blocking antibody" through competitive inhibition of IgE attachment and an inhibitory influence on IgE-induced allergic reaction (Platts-Mills *et al.*, 2021). Legrand (2016) also mentioned C4A immune-modulatory effect, which has a specific influence in tolerance induction at early developmental stages in immunity.

Besides that, the study by Lupinek *et al.* (2019) shows that mothers form specific IgG antibodies into ubiquitous allergens during pregnancy, and they transfer these to their babies through the placenta and in breast milk. The maternal IgG antibodies protect the babies by inhibiting allergens from provoking allergic reactions. As an example, allergens like timothy grass pollen (Phl p 1), birch pollen (Bet v 1), cat allergen (Fel d 1), wasp venom (Ves v 5), and common food allergens like peanut (Ara h 1), cow's milk (Bos d 5), egg (Gal d 1), and wheat (Tri a 36) were included in the study. If mothers have high levels of IgG antibodies against these allergens, their children are less likely to become allergically sensitized (IgE response) to the same allergens by the time they are five years old. The protective effect is because maternal IgG can neutralize allergens or remove them before they can trigger an allergic reaction. The babies start making their own IgG antibodies around six months, but the maternal IgG shields them during the initial six months when their immune system is not developed yet. In the KEGG pathway Fc gamma R-mediated phagocytosis panel (hsa04666), it is demonstrated, where attachment of IgG-allergen complexes through Fcγ receptors on leukocytes with subsequent clearance through phagocytosis inhibits inflammation. In the Antigen Processing and Presentation panel panel (hsa04612), IgG determines antigen presentation toward a tolerogenic direction, with a hope for immune tolerance. Also, in the Complement and Coagulation Cascades panel panel (hsa04610), IgG activates the classical complement system for elimination of immune complexes. In the processes, there are descriptions where high maternal IgG level, especially above 30 ISU-G, results in protection for allergic sensitization in childhood. It is also an ongoing protection demonstrated by Ballard and Morrow (2013), where there were descriptions for how maternal IgG gives early protection for infection. Rajani *et al.* (2018) also showed a long-term function for maternal IgG in infants, with a lower

risk for allergic disease like eczema, also for asthma. In total, there are clear results demonstrating a clear proof for how maternal IgG, both through passive transfer, also through molecular immunity mechanisms, plays a crucial role in providing for the infant's first environment for immunity, with protection for developing an allergic condition.

Lastly, Manti *et al.* (2016) provided evidence that breast milk anti-inflammatory cytokine interleukin-10 (IL-10) plays a key function in controlling allergic reactions in infants. Breastfeeding CMPA babies had higher serum IL-10 levels than non-breastfeeding CMPA babies, and these higher levels were correlated with milder symptoms of allergy and lower SCORAD eczema scores. IL-10 was also significantly and negatively correlated with total IgE levels, indicating IL-10 plays a suppressive function against allergic inflammation. Such molecular mechanisms are supported in the directionality of results through KEGG pathway analysis: in the cytokine-cytokine receptor interaction pathway (hsa04060), IL-10 binds through the IL10R receptor, triggering downstream anti-inflammatory signaling; in the Jak-STAT signaling pathway (hsa04630), IL-10 activates STAT3, the latter triggering anti-immunestimulatory gene expression; in the T cell receptor signaling pathway (hsa04660), IL-10 activates regulatory T-cell production for preservation of immunological tolerance. Such molecular mechanisms are consistent with Legrand (2016), who highlighted IL-10's action in suppressing Th2 responses and crucial in maintaining Th1/Th2 balance towards prevention of hypersensitivity. In aggregate, these findings solidify IL-10 as a critical cytokine for protection against risk for allergy/immunity homeostasis during early life.

Although numerous studies on human milk proteomes and their potential to prevent allergy have reported promising findings, it is important to consider the confounding variables that could influence the findings of these studies. Donald *et al.* (2024) suggested that maternal age, mode of delivery, and early antibiotic exposure are confounding variables likely to influence the amount and quality of transferred secretory immunoglobulin A (SIgA) via breastfeeding. Jiang *et al.* (2021) accounted for geographical variables, maternal genetic characteristics, and secretor status and found that a strong association between Factor 2 proteome and reduced allergy risk was attenuated upon adjustment for geographical confounding variables. This suggests that the environment and lifestyle variables could be significant contributors. At the same time, Lupinek *et al.* (2019) suggested that exposure history of a mother, immune status, and breastfeeding exclusivity might determine the amount and specificity of transferred infant IgG, crucial to modulating the infant's immune response. Finally, Manti *et al.* (2016) recognized that type of feeding, allergen exposure duration, and infant immunity variability could influence IL-10 production, which in turn influences the expression of allergic diseases in infants. These studies collectively underscore the need for future research to control for these variables to better isolate the effects of human milk proteomes in allergy prevention.

CONCLUSION

In conclusion, this study highlights the role of human milk proteomes in preventing allergen sensitivity in infants. Seven proteomes were identified in this review which are secretory immunoglobulin A, IgG, vitronectin, CD81, Complement C4A, Immunoglobulin γ4, and IL-10. The bioactive components in human milk contribute significantly in building the immune system through passive immunity, immunomodulation, and stimulation of gut health. These bioactive molecules inhibit allergic sensitization, protect against infections, invoke a synergistic immune response and thereby reduce the likelihood of developing allergic diseases like asthma, eczema, and allergic rhinitis. While the current evidence highlights the protective role of human milk proteomes in reducing allergy risk in infants, further research is needed to uncover the underlying mechanisms and long-term benefits. This includes methodologically robust studies that control for key variables such as maternal diet, genetics, and environment, while applying longitudinal designs and standardized protocols. A broader, multi-omics approach may also provide a more integrated understanding of how various milk components work together to promote immune tolerance. These efforts can ultimately inform targeted strategies for allergy prevention and early-life health promotion.

ACKNOWLEDGEMENT

Appreciations are extended to all involved in the making of this review article and the Department of Biomedical Science of IIUM for the support.

REFERENCES

Arksey, H., & O'Malley, L. (2005).Scoping studies: towards a methodological framework. International Journal of Social Research Methodology, 19-32.

ASCIA. (2020). *ASCIA guidelines - infant feeding and allergy prevention*. Australasian Society of Clinical Immunology and Allergy (ASCIA). https://www.allergy.org.au/hp/papers/infant-feeding-

and-allergy-prevention

Ballard, O., & Morrow, A. L. (2013). Human milk composition: Nutrients and bioactive factors. *Pediatric Clinics of North America*, *60*(1), 49–74. https://doi.org/10.1016/j.pcl.2012.10.002

Donald, K., Serapio-Palacios, A., Gerbec, Z., Bozorgmehr, T., Holani, R., Cruz, A. R., Schnupf, P., & Finlay, B. B. (2024). Secretory IgA in breast milk protects against asthma through modulation of the gut microbiota. *Cell Reports*, *43*(10), 114835–114835. https://doi.org/10.1016/j.celrep.2024.114835

Gutowska-Ślesik, J., Samoliński, B., & Krzych-Fałta, E. (2023). The increase in allergic conditions based on a review of literature. *Advances in Dermatology and Allergology/Postępy Dermatologii I Alergologii*, *40*(1), 1. https://doi.org/10.5114/ada.2022.119009

Jiang, S., Pan, J., Li, Y., Ju, M., Zhang, W., Lu, J., Lv, J., & Li, K. (2021). Comprehensive human milk patterns are related to infant growth and allergy in the CHMP study. *Molecular Nutrition & Food Research*, *65*(17). https://doi.org/10.1002/mnfr.202100011

Kull, I., Melen, E., Alm, J., Hallberg, J., Svartengren, M., van Hage, M., Pershagen, G., Wickman, M., & Bergström, A. (2010). Breast-feeding in relation to asthma, lung function, and sensitization in young schoolchildren. *Journal of Allergy and Clinical Immunology*, *125*(5), 1013–1019. https://doi.org/10.1016/j.jaci.2010.01.051

Legrand, D. (2016). Overview of lactoferrin as a natural immune modulator. *The Journal of Pediatrics*, *173*, S10–S15. https://doi.org/10.1016/j.jpeds.2016.02.071

Lupinek, C., Hochwallner, H., Johansson, C., Mie, A., Rigler, E., Scheynius, A., Alm, J., & Valenta, R. (2019). Maternal allergen-specific IgG might protect the child against allergic sensitization. *Journal of Allergy and Clinical Immunology*, *144*(2), 536–548. https://doi.org/10.1016/j.jaci.2018.11.051

Manti, S., Lougaris, V., Cuppari, C., Tardino, L., Dipasquale, V., Arrigo, T., Salpietro, C., & Leonardi, S. (2017). Breastfeeding and IL-10 levels in children affected by cow's milk protein allergy: A restrospective study. *Immunobiology*, *222*(2), 358–362. https://doi.org/10.1016/j.imbio.2016.09.003

Oh, J., Kim, S., Kim, M. S., Abate, Y. H., Abd ElHafeez, S., Abdelkader, A., Abdi, P., Abdulah, D. M., Aboagye, R. G., Abolhassani, H., Abtahi, D., Abualruz, H., Abu-Gharbieh, E., Aburuz, S., Adane, M. M., Addo, I. Y., Adeleke, O. T., Aden, B., Adnani, Q. E. S., & Adra, S. (2025). Global, regional, and national burden of asthma and atopic dermatitis, 1990–2021, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. *The Lancet Respiratory Medicine*, *13*(5). https://doi.org/10.1016/s2213-2600(25)00003-7

Platts-Mills, T. A. E., Keshavarz, B., Wilson, J. M., Li, R., Heymann, P. W., Gold, D. R., McGowan, E. C., & Erwin, E. A. (2021). An overview of the relevance of IgG4 antibodies in allergic disease with a focus on food allergens. *Children*, *8*(5), 418. https://doi.org/10.3390/children8050418

PRISMA. (2020). *PRISMA 2020 flow diagram*. PRISMA. https://www.prisma-statement.org/prisma-2020-flow-diagram

Rahman, T., Sarwar, P. F., Potter, C., Comstock, S. S., & Vanja Klepac-Ceraj. (2023). Role of human milk oligosaccharide metabolizing bacteria in the development of atopic dermatitis/eczema. *Frontiers in Pediatrics*, *11*. https://doi.org/10.3389/fped.2023.1090048

Rajani, P. S., Seppo, A. E., & Järvinen, K. M. (2018). Immunologically active components in human milk and development of atopic disease, with emphasis on food allergy, in the pediatric population. *Frontiers in Pediatrics*, *6*. https://doi.org/10.3389/fped.2018.00218

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., & Lewin, S. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Zhu, J., & Dingess, K. A. (2019). The functional power of the human milk proteome. *Nutrients*, *11*(8), 1834. https://doi.org/10.3390/nu11081834

ORIGINAL ARTICLE

Antioxidant-Treated Sperm Enhances Fertilization and Early Embryogenesis in Intracytoplasmic Sperm Injection (ICSI) Cycles

Azantee Yazmie Abdul Wahab[1,*], Sardiana Sarmidi[2]

[1]Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2]Reproductive Medicine Unit, Department of Obstetrics and Gynaecology, Hospital Tunku Azizah, Jalan Pahang, 50586 Kuala Lumpur, Malaysia

Keywords:
antioxidant supplement; oxidative stress; early embryogenesis; intracytoplasmic sperm injection (ICSI); fertilization

ABSTRACT

Background: Oxidative stress (OS) is a major factor influencing sperm quality and has been implicated in 30–80% of male factor infertility cases, often resulting in compromised fertilization capacity and diminished embryo quality in assisted reproductive technology (ART). While antioxidant supplementation has demonstrated beneficial effects on sperm parameters, particularly in European and American populations, evidence from Asian cohorts remains scarce. Moreover, the role of antioxidants in modulating early embryogenesis following Intracytoplasmic Sperm Injection (ICSI) is not yet well established. This study aimed to evaluate the association between antioxidant supplementation, specifically PROfertil®**,** and early embryogenesis outcomes in Malaysian couples affected by Oligoasthenoteratozoospermia (OAT). **Methods:** A single-center retrospective analysis was conducted on 195 female partners of OAT-affected males who underwent ART treatment between 2016 and 2021 at the Reproductive Medicine Unit, Hospital Tunku Azizah. Male partners in the treatment group received PROfertil® (two capsules daily for three months) prior to ART, while those in the control group received no supplementation. A total of 783 oocytes were retrieved from the treatment group and 796 from the control group. Statistical anaylsis were done on the early embryogenesis including fertilization, cleavage and blastocyst rates as well as embryo quality. **Results:** Comparative analysis using Kruskal-Wallis and Chi-Square tests revealed significant differences. Demographic analysis showed statistical differences in husband's age and the etiology of infertility ($p<0.05$). The treatment group exhibited higher percentages of fertilization, cleavage on day 2, good cleavage on day 3, as well as quality embryo grading, including implantation potential on day 3 ($p < 0.05$). However, statistical analysis showed significant differences on day 5 in the treatment group compared to control group, with fewer good-quality blastocysts observed ($p = 0.03$). **Conclusion:** The findings suggest that antioxidant supplementation with PROfertil® improves the sperm quality and enhancing early embryogenesis outcomes in couples struggling with male factor infertility, particularly OAT-related cases. Further large-scale prospective studies are warranted to validate these results and to establish the clinical relevance of antioxidant supplementation in ART practice.

INTRODUCTION

A delicate balance of reactive oxygen species (ROS) is essential for proper sperm function. However, excessive ROS levels can detrimentally affect sperm quality and fertilization capacity. Oxidative stress (OS) has garnered significant attention due to ROS and their byproducts damaging DNA, lipids, proteins, and enzymatic systems, leading to cell death and deteriorating semen parameters associated with male infertility (Dutta et al., 2019). Antioxidant supplementation protects sperm from oxidative stress by neutralizing harmful free radicals, which helps to preserve sperm's structural integrity, DNA integrity, motility, and overall function. Commonly used antioxidants include vitamin E, vitamin C, carnitines, N-acetyl cysteine, CoQ10, zinc, selenium, folic acid, and lycopene, either as monotherapy or in combination.

A comprehensive review by Majzoub and Agarwal (2018) highlighted the positive impact of antioxidants on male fertility across 26 studies, covering basic semen parameters, advanced sperm function, and outcomes of assisted reproductive technology (ART). A Cochrane review by Smits et al. (2019) reported that antioxidants were effective in enhancing clinical pregnancy rates compared to placebo or no treatment. This conclusion was based on 105 clinical pregnancies from 786 couples across 11 small studies (Busetto et al., 2018; Barekat et al., 2016; Attallah et al., 2013). A study by Tremellen et al. (2007) stated that antioxidant supplement; Menevit (Bayer, Sydney, Australia), which contained Lycopene 6 mg, Vitamin E 400 IU, Vitamin C 100 mg, Zinc 25 mg, Selenium 26 µg, Folate 0.5 mg, and Garlic 1000 mg showed no significant effect on embryo quality and fertilization rates in IVF-ICSI cycles, but did demonstrate an improvement in pregnancy rates. These studies shown that antioxidant supplementation may positively influence pregnancy rates in assisted reproductive contexts.

* *Corresponding author.*
E-mail address: yazmie@iium.edu.my

PROfertil® is a formulated antioxidant supplement developed to enhance sperm quality. It comprises eight essential micronutrients: L-carnitine (440 mg), L-arginine (250 mg), zinc (40 mg), vitamin E (120 mg), glutathione (80 mg), selenium (60 µg), coenzyme Q10 (15 mg), and folic acid (800 µg). Previous clinical studies have reported that PROfertil® supplementation significantly improves semen parameters, including concentration, motility, and morphology (Lipovac et al., 2016; Lipovac et al., 2014; Imhof et al., 2012; Schauer, 2011). Reported spontaneous pregnancy rates following six months of supplementation ranged from 25.8% to 27.8% (Imhof et al., 2012; Lipovac et al., 2021). Likewise, Schauer et al. (2011) demonstrated that PROfertil® enhanced sperm morphology and was associated with a 41.18% spontaneous pregnancy rate.

To date, no study has specifically investigated the effects of PROfertil®-treated sperm on embryo development following intracytoplasmic sperm injection (ICSI) cycles. Therefore, the present study aimed to explore the role of PROfertil® in supporting embryonic development after ICSI. Specifically, we evaluated the impact of PROfertil®-treated sperm on early embryogenesis, including fertilization, cleavage, and blastocyst formation rates, as well as overall embryo quality. Following the documented improvements in sperm quality, the subsequent phase of this study involved performing ICSI using oocytes retrieved from the female partners of the same couples undergoing assisted reproductive treatment (ART).

MATERIALS AND METHODS

Study recruitment

This retrospective study analyzed 1,579 oocytes collected from 195 female partners who underwent ICSI at the Reproductive Medicine Unit, Hospital Tunku Azizah (Hospital Wanita dan Kanak-Kanak Kuala Lumpur) between 1 January 2016 and 19 April 2021. Their corresponding male partners, all diagnosed with oligoasthenoteratozoospermia (OAT), were divided into two groups: (i) the treatment group, in which men received two daily capsules of the antioxidant supplement PROfertil® (LenusPharma, Vienna, Austria) for three months as prescribed by a clinician, and (ii) the control group, in which men received no micronutrient supplementation.

The inclusion criteria for male partners were subfertility for at least one year, age between 20 and 50 years, and at least one prior abnormal semen analysis consistent with OAT, defined as low sperm concentration (<15 million/mL), reduced motility (<32%), and normal morphology <4% (Cooper et al., 2010; WHO, 2010). Exclusion criteria included infertile male partners with azoospermia, aspermia, varicocele, or recent urogenital infections. Patients with diabetes, inflammatory disease, chronic kidney or renal disease, or those receiving hormone therapy or antioxidant supplements were also excluded from the study. The inclusion criteria for female partners were couples who underwent an ICSI cycle with an infertile male partner and had been trying to conceive for at least one year. The exclusion criteria included female partners with diminished ovarian reserve or endometriosis.

Ethical approval

Ethical approval was obtained from the Medical Research and Ethics Committee (MREC) and the Clinical Research Centre (CRC), Hospital Tunku Azizah (NMRR-20-2975-57354 [IIR]). All subject information was handled in accordance with the approved protocol, including justification for the inclusion criteria.

Sperm preparation for ICSI

An aliquot of 5-10 µL from the same sperm suspension was used for sperm analysis. Then, 0.5-2 ml of semen sample was prepared using sperm preparation (swim up, density gradient or simple wash) based on the quality of sperm such as concentration and motility. The final diluted sperm was placed at room temperature before use for ICSI.

Oocyte retrieval

The female partners underwent controlled ovarian stimulation with injectable follicle-stimulating hormone (FSH) for approximately 8–12 days until multiple mature follicles developed. In addition to FSH, gonadotropin-releasing hormone (GnRH) agonists or antagonists were administered to prevent premature ovulation. Follicular development was monitored by serial transvaginal ultrasound, and once a dominant follicle measuring ≥18 mm was observed, human chorionic gonadotropin (hCG) was administered to trigger final oocyte maturation. Oocyte retrieval was performed 34–36 hours after hCG administration via transvaginal ultrasound-guided aspiration. The collected follicular fluid was immediately examined by embryologists to identify and isolate oocytes. A total of 783 oocytes were retrieved from the treatment group and 796 from the control group.

Intracytoplasmic sperm injection (ICSI) technique

Mature oocytes at the metaphase II stage were selected for ICSI. The oocytes were stabilized with a holding pipette, and a single sperm was carefully injected into the cytoplasm using a micromanipulation system

(Eppendorf® Micromanipulator, US).

Embryo culture

Injected and fertilized oocytes were incubated in Global total LP medium (LifeGlobal, Belgium) at 37°C with 6% CO2 until day 5 in the same medium and conditions.

Embryo analysis

Fertilization check

The oocytes were examined under an inverted microscope 16–18 hours after injection. Fertilization was confirmed by the presence of two pronuclei (2PN). The fertilization rate (FR) was calculated as the percentage of fertilized oocytes relative to the number of injected oocytes.

Assessment of embryo quality

Embryo development was assessed daily from day 2 to day 5, except on day 4. Observations were performed under a microscope at predetermined time points, as shown in **Table 1**. Embryo cleavage quality was evaluated and recorded according to the criteria in **Table 2**. The grading system was based on the number of cells (blastomeres), cell size and shape, and the percentage of fragmentation (Balaban et al., 2011). On day 5, embryos were assessed for progression to the blastocyst stage and graded using Gardner's classification, as presented in **Table 3**, which includes blastocoele expansion, inner cell mass (ICM) size and trophectoderm (TE) cohesion. Blastocysts graded 3–5AA were considered excellent, those graded 3–4AB/BA/BB were considered good, those graded 3–4BC/CB/CA/AC were considered poor, and those graded 2–4CC/EB were considered very poor.

Embryo transfer

Two types of embryo transfer were performed: fresh embryo transfer (ET) and frozen embryo transfer (FET). ET was carried out when the embryo reached day 2–5, depending on embryo quality and the condition of the uterine cavity. Patients who were not suitable for fresh ET due to factors such as a high risk of ovarian hyperstimulation syndrome (OHSS) or an unfavorable uterine environment were advised to undergo FET. Embryo transfer, either fresh or frozen, was performed on day 2, day 3, or day 5.

Clinical outcomes following ICSI

Serum β-hCG was measured two weeks after ET. A β-hCG level >50 IU/L was considered positive for biochemical pregnancy, and clinical pregnancy was confirmed at the subsequent transvaginal scan (TVS) by the presence of a gestational sac or fetal heartbeat (fetal pole). Live birth was defined as the delivery of a baby born alive. Pregnancy, miscarriage, and live birth rates were recorded.

Statistical Analysis

Data on demographic and baseline characteristics were expressed as numbers and percentages and analyzed using the Kruskal–Wallis test. Early embryogenesis parameters (fertilization, cleavage, blastocyst rates, and embryo quality) were also presented as numbers and percentages and analyzed using Chi-square tests. A p-value <0.05 was considered statistically significant.

Table 1: Timing of observation of fertilized oocytes and embryos, and expected stage of development at each time point. Adapted from Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology (2011).

Type of observation	Timing (hours post insemination)	Expected stage of development
Fertilization check	17±1	Pronuclear stage
Syngamy	23±1	Expect 50 % to be in syngamy (up to 20% may be at the 2-cell stage)
Early Cleavage Stage	26±1 (ICSI) 28±1 (IVF)	2 cell stage
Day-2 embryo assessment	44±1	4 cell stage
Day-3 embryo assessment	68±1	8 cell stage
Day-4 embryo assessment	92±2	Morula
Day-5 embryo assessment	116±2	Blastocyst

Table 2: Embryo cleavage scoring system. Adapted from Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology (2011)

Grade	Rating	Description
1	Good	• <10% fragmentation • Stage-specific cell size • No multinucleation
2	Fair	• 10– 25% fragmentation • Stage-specific cell size for the majority of cells • No evidence of multinucleation
3	Poor	• Severe fragmentation (>25%) • Cell size not stage specific • Evidence of multinucleation

Table 3: Blastocyst scoring system based on Gardner's classification. Adapted from Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology (2011).

Size and expansion	
Classification	**Description**
1	Early blastocyst: the blastocyst is less than half the volume of the embryo.
2	Blastocyst: the blastocele is greater than or equal to half of the volume of the embryo.
3	Full blastocyst: the blastocele completely fills the embryo.
4	Expanded blastocyst: the blastocele volume is larger than that of the early embryo and the zona pellucida is thinning.
5	Hatching blastocyst: the trophectoderm has started to herniate through the zona pellucida.
6	Hatched blastocyst: the blastocyst has completely escaped from the zona pellucida.
Inner cell mass (ICM)	
Classification	**Description**
A	Tightly packed, many cells
B	Loosely grouped, several cells
C	Very few cells
Trophectoderm (TE)	
Classification	**Description**
A	Many cells forming a tightly knit epithelium.
B	Few cells.
C	Very few cells forming a loose epithelium.

Notes: The scoring system for blastocysts is a combination of the size and expansion and of the grade of the ICM and of the TE (e.g., an expanded blastocyst with a good ICM and a good TE would be scored as 4AA).

RESULTS

Patient Demographic and Baseline Characteristics.

A total of 200 subfertile men whose female partners underwent ICSI and had at least one recent abnormal semen analysis at the Reproductive Medicine Unit, Hospital Tunku Azizah, were recruited for this study. Four participants withdrew, leaving 92 infertile men who completed three months of antioxidant supplementation (PROfertil®) and 103 men who did not take any supplements as the control group.

The mean (SD) age of men in the PROfertil® group was 37.04 (6.83), while the control group had a mean age of 35.30 (5.40). **Table 4** summarizes the demographic and baseline characteristics, showing significant differences (p<0.05) in the husbands' age and infertility etiology, while other parameters were similar between groups.

Interestingly, the treatment group was slightly younger, suggesting that younger men may be more proactive in seeking fertility treatment. Additionally, a higher proportion of men (70%) with male factor infertility received PROfertil® compared to those with combined infertility factors (30%) (p=0.02). This indicates that male partners with identified male factor infertility were more likely to undergo antioxidant therapy to improve sperm quality.

Embryo Development Rate

Fertilization and abnormal rate

The fertilization rate was calculated as the number of fertilized oocytes divided by the total number of oocytes observed per group after 16-18 hours of insemination through ICSI. Based on **Table 5**, there was a significant difference between both groups *(p*=0.003).

Abnormal fertilization rate was calculated as the number of fertilised oocytes in which consist of three or more pronuclei (>3PN) or only one pronuclei (1PN) and divided by the total number of oocytes observed per group after 16-18 hours of insemination through ICSI. Based on **Table 5,** there was no significant difference in the abnormal fertilisation rate between the PROfertil® and control groups (p=0.48).

Cleavage Rate

Cleavage rates were calculated by determining the number of embryos at the 4-cell stage on day 2 (44 ± 1 hours after ICSI) and the number of embryos at the 8-cell stage on day 3 (68 ± 1 hours after ICSI), divided by the total number of fertilized oocytes. Good cleavage rates were calculated based on the number of embryos graded as grade 1 or 2 on day 2 and day 3, according to the embryo cleavage scoring system by the Alpha Scientists in Reproductive Medicine and the ESHRE Special Interest Group of Embryology (2011).

Table 4: Demographic and baseline characteristics of the population by treatment group (i.e. PROfertil®) and control group**.**

CHARACTERISTICS	CONTROL (n=103)	PROfertil® (n=92)	*p*-VALUE
Age[a]			
Wife (mean, SD)	34.07 (4.45)	33.05 (3.67)	0.08
Husband (mean, SD)	37.04 (6.83)	35.30 (5.40)	**0.02***
Ethnicity[b]			
Malay (n,%)	64 (62)	59 (64)	0.96
Chinese (n,%)	15 (14.6)	13 (14)	
Indian (n,%)	24 (23.3)	20 (22)	
BMI of wife[b]			
Underweight (n,%)	2(2)	4(4)	0.65
Normal (n,%)	48 (47)	47(51)	
Overweight (n,%)	45 (44)	34 (37)	
Obese (n,%)	8 (8)	7 (8)	
BMI of husband[b]			
Underweight (n,%)	1 (1)	1 (1)	0.11
Normal (n,%)	34 (33)	19 (21)	
Overweight (n,%)	49 (48)	43 (47)	
Obese (n,%)	19 (18)	29 (32)	
Smoking[b]			
Yes (n,%)	12 (12)	16 (17)	0.25
No (n,%)	91 (88)	76 (83)	
Etiology of infertility[b]			
Male (n,%)	56 (54)	64 (70)	**0.02***
Combined (n,%)	47(46)	28 (30)	

Notes: *Significant differences as compared between different groups at $p<0.05$. [a]Data are expressed as mean (SD) and analyzed using Kruskal-Wallis test. [b]Data are expressed as frequency (%) and analyzed using the Chi-Square test. Independent variables were presented as demographic and baseline characteristics (i.e age of wife/husband, ethnicity, etiology of infertility, husband/wife BMI and primary/secondary fertility). Dependent variables were presented as the control or treatment (i.e outcomes).

As shown in **Table 5**, there was a significant difference in cleavage rate on day 2 (p = 0.004), with the treatment group showing a slightly higher percentage compared to the control group (99.8% vs. 98.2%). On day 3, there was also a significant difference in the percentage of good cleavage, with the PROfertil® group showing a higher rate compared to the control group (83.7% vs. 78.0%, p = 0.012). However, no significant differences were observed in good cleavage on day 2 (p = 0.15) or in overall cleavage on day 3 (p = 0.66) between the two groups, as presented in **Table 5**.

Multinucleated blastomeres rate (MNB)

The MNB rate was calculated as the number of embryos that present two, three, or more nuclei in their blastomeres and divided by the number total of fertilized oocytes observed per group on cleavage stage. Based on **Table 5,** there was no significant difference in the rate of MNB on cleavage stage between the PROfertil® and control groups (p=0.32).

Table 5: Comparison of the embryo development rate between the treatment (PROfertil®) and control group.

Fertilization rate	**Control group (n=796)**	**Treatment group (i.e.PROfertil®) (n=783)**	**Relative ratio**	***p* value**
Fertilisation noted 16-18H	568 (71%)	609 (77.8%)	1.09 (95%CI 1.03, 1.15)	**0.003***
Abnormal fertilization	25 (3.1%)	20 (2.6%)	0.81 (95%CI 0.46, 1.45)	0.48
Cleavage rate	**Control group (n=568)**	**Treatment group (i.e.PROfertil®) (n=609)**	**Relative ratio**	***p-value***
Cleavage on Day 2	558 (98.2%)	608 (99.8%)	1.02 (95%CI, 1.0 to 1.03)	**0.004***
Good cleavage Day 2	497 (87.5%)	549 (90.1%)	1.03 (95%CI, 0.99 to 1.07)	0.15
Cleavage on Day 3	563 (99.1%)	605 (99.3%)	1.00 (95%CI 0.99 to 1.01)	0.66
Good cleavage Day 3	443 (78.0%)	510 (83.7%)	1.07 (95%CI, 1.02 to 1.14)	**0.012***
Multinucleated blastomeres (MNB)	7 (1.2%)	12 (2.0%)	1.60 (95%CI 0.63, 4.03)	0.32
Blastocyst rate	**Control group (ᶻ=208)**	**Treatment group (i.e.PROfertil®) (ᶻ=222)**	**Relative ratio**	***p-value***
[1]Blastocyst	98 (47.1%)	85 (38.3%)	0.81 (95%CI 0.65 to 1.01)	0.06
[1]Good Blastocyst	35 (16.8%)	22 (9.9%)	0.59 (95%CI, 0.36 to 0.97)	**0.03***

Notes: Data was presented as numbers (%) and analysed using Chi-Square test; n is the number of injected oocytes following ICSI; ᶻ is the number of embryos pushed to the blastocyst stage; [1]Total of blastocysts formed out of the number of embryos pushed to blastocyst stage; *Significant difference between PROfertil® and control groups $p<0.05$.

Blastocyst Rate

The blastocyst rate was calculated by dividing the number of blastocysts formed on day 5 of culture (116 ± 2 hours after ICSI) by the total number of embryos cultured until day 5 in each group. Good blastocysts were defined as those graded 3–5AA (excellent) or 3–5AB/BA/BB (good). As shown in **Table 5**, the percentage of blastocysts formed was slightly lower in the treatment group compared to the control group (38.3% vs. 47.1%), as was the percentage of good blastocysts (9.9% vs. 16.8%). Statistical analysis revealed a significant difference in the good blastocyst rate between the two groups (p = 0.03).

Grade of quality embryos and blastocyst

Cleavage stage

Embryo cleavage was assessed by grading based on the appearance of blastomeres, cytoplasmic defects, and the degree of fragmentation on day 2 or day 3 of culture. As shown in **Table 6**, the treatment group had a slightly higher number of grade 1 (good) and grade 2 (fair) embryos, and a slightly lower number of grade 3 (poor) embryos on day 2 compared to the control group; however, the difference was not statistically significant. Additionally, the PROfertil® group showed a slightly higher percentage of embryos graded as good and fair for potential implantation compared to the control group (90.3% vs. 89.1%), though this difference was also not significant.

In contrast, **Table 6** demonstrated a significant difference in embryo quality on day 3 (p = 0.039). The treatment group had a higher percentage of embryos graded as good and fair for potential implantation compared to the control group (p = 0.013). The proportion of poor-quality embryos was lower in the treatment group than in the control group on both day 2 and day 3, with a statistically significant difference observed only on day 3 (p = 0.013)

Blastocyst stage

In **Table 6**, no significant difference was observed on the fifth day between the treatment and control groups in terms of the distribution of excellent, good, poor, and very poor blastocysts. Additionally, there was also no significant difference in the percentage of embryos graded as good and poor quality for potential implantation between the two groups.

Table 6: The comparison of the quality of the embryos and blastocyst between the control and treatment groups (PROfertil®)

Grade of quality embryos (Day 2)	**Control group (n=558)**	**Treatment group (n=608)**	**Relative ratio**	***p-value***
Good	61 (10.9%)	65 (10.7%)	-	0.77
Fair	436 (78.2%)	484 (79.6%)	-	
Poor	61 (10.9%)	59 (9.7%)	-	
Considerable for implantation (Good and fair embryo)	497 (89.1%)	549 (90.3%)	1.014 (95% CI, 0.975 to 1.054)	0.49
Not considerable for implantation (Poor embryos)	61 (10.9%)	59 (9.7%)	0.888 (95%CI, 0.632 to 1.246)	
Grade of quality embryos (Day 3)	**Control group (n=563)**	**Treatment group (n=605)**	**Relative ratio**	***p-value***
Good	48 (8.5%)	49 (8.1%)	-	**0.039***
Fair	395 (70.2%)	461 (76.2%)	-	
Poor	120 (21.3%)	95 (15.7%)	-	
Considerable for implantation (Good and fair graded embryo)	443 (78.7%)	510 (84.3%)	1.071 (95% CI, 1.014,1.132)	**0.013***
Not considerable for implantation (Poor embryos)	120 (21.3%)	95 (15.7%)	0.737 (95%CI, 0.578 to 0.940)	
Grade of quality blastocyst (Day 5)	**Control group (ᶻ=98)**	**Treatment group (ᶻ=85)**	**Relative ratio**	***p-value***
[1]Excellent	9 (9.2%)	11 (12.9%)	-	0.12
[1]Good	26 (26.5%)	11 (12.9%)	-	
[1]Poor	18 (18.4%)	15 (17.6%)	-	
[1]Very poor	45 (45.9%)	48 (56.5%)	-	
[1]Considerable for implantation (Excellent and good blastocyst)	35 (35.7%)	22(25.9%)	0.72 (95%CI, 0.46, 1.13)	0.15
Not considerable for implantation (Poor and very poor blastocyst)	63 (74.1%)	63 (64.3%)	1.15 (95%CI 0.95 to 1.40)	

Notes: Data were presented as numbers (%) and analysed using Chi-Square test; n is the number of fertilised oocytes; ᶻ is the number of embryos pushed to the blastocyst stage; [1]Total of blastocysts formed out of the number of embryos pushed to blastocyst stage; *Significant difference between PROfertil® and control groups $p<0.05$

DISCUSSION

The mean age of participants in the treatment group was slightly higher than that of the control group, 37.04 (6.83) versus 35.30 (5.40), respectively, with a statistically significant difference ($p < 0.05$). This finding aligns with demographic trends reported in other fertility studies, where age differences among male partners are often associated with variations in reproductive health–seeking behavior. Although sperm quality is generally believed to decline with advancing age, recent evidence by Pino *et al.* (2020) suggests that a significant reduction in sperm parameters may not occur until after the age of 50. Therefore, the age difference observed between the two

groups in the present study is unlikely to have had a substantial impact on sperm quality outcomes.

Our findings also revealed a significant difference in the etiology of infertility between groups. A larger proportion of men in the treatment group (70%) were diagnosed with male factor infertility compared to those classified under combined infertility factors (30%) (*p*=0.02). This pattern suggests that men with a confirmed male factor diagnosis are more inclined to pursue antioxidant supplementation, such as PROfertil®, in an effort to enhance sperm quality and improve their chances of conception through assisted reproductive technologies.

Various factors contribute to the complex process of fertilization, with sperm quality playing a crucial role. Sperm quality including concentration, motility, and morpholog is critical for successful fertilization (Vogiatzi et al., 2022; Bartolacci et al., 2018; Zhu et al., 2019). As shown in Table 5, fertilization rates were significantly higher in the PROfertil® group compared to the control group (p = 0.003), suggesting that supplementation may enhance sperm–oocyte interaction. This finding is consistent with Lipovac et al. (2014), who reported improved sperm-hyaluronan binding assay (SHBA) values following PROfertil® use, reflecting better fertilization potential. Similarly, Rashki Ghaleno et al. (2016) noted that SHBA outcomes depend on sperm concentration and motility. Supporting evidence from Gharagozloo et al. (2016) demonstrated that supplementation with antioxidants present in PROfertil® (vitamin C, vitamin E, carnitine, folic acid, lycopene, selenium, and zinc) in a mouse model led to a twofold increase in fertilization rates compared to non-supplemented animals (73.7% vs. 35.2%). Collectively, these findings highlight the potential benefits of PROfertil®, likely driven by its antioxidant components, in improving fertilization outcomes.

In our finding revealed no significant difference in terms of abnormal fertilization in both groups (p=0.48) as shown in Table 5. Some previous studies have looked into the effect of antioxidant oral supplements on IVF/ICSI cycle clinical outcomes (Scaruffi et al., 2021; Gambera et al., 2019; Karosi et al., 2017; Rago et al., 2017; Tremellen et al., 2007) but they did not report the abnormal fertilization result on their findings. However, Pantos et al. (2021) found that there was no connection between abnormalities in semen analysis and the rate of 1PN or 3PN formation.

The quality of cleavage embryos is the key determinant of success and reflects their potential to develop into robust blastocysts. As shown in Table 5, no significant difference was observed in cleavage rates on day 3 between the PROfertil® and control groups (p > 0.05). This is consistent with Greco et al. (2005) and Tremellen et al. (2021), who also reported no effect of antioxidant supplementation on day 2–3 cleavage rates. However, significant improvements were found in our study for good cleavage on day 3 (83.7% vs. 78.0%; p = 0.012) and overall cleavage on day 2 (99.8% vs. 98.2%; p < 0.05) in the PROfertil® group. These findings suggest that 3 months of PROfertil® treatment enhances embryo numbers post-ICSI. Supporting evidence comes from Gambera et al. (2019) and Scaruffi et al. (2021), who also observed increased cleavage rates and embryo numbers with different antioxidant supplement combinations.

The Alpha Scientists in Reproductive Medicine and the ESHRE Special Interest Group of Embryology (2011) defined multinucleated blastomeres (MNB) as blastomeres containing two or more interphase nuclei. MNB has been associated with poor embryo growth, aneuploidy, chromosomal abnormalities, and unfavorable IVF outcomes (Balakier et al., 2016; Desai et al., 2016). In our study, no significant difference in MNB incidence was observed between the PROfertil® and control groups (p = 0.32). Desch et al. (2017) reported lower implantation (27.7%) and live birth rates (22.7%) in embryos with MNB compared to those without (33.4% and 29.8%). Some MNB embryos, however, can self-correct during early cleavage, developing into euploid blastocysts and healthy offspring (Egashira et al., 2015; Zhu et al., 2014). Studies on antioxidant supplements and ART outcomes have not specifically reported on MNB (Scaruffi et al., 2021; Gambera et al., 2019; Karosi et al., 2017; Rago et al., 2017; Tremellen et al., 2007).

Fragmentation, cell debris, cell number, size, and symmetry are key markers of embryo quality in ART. In our study, the PROfertil® group showed higher favorable cleavage on day 3 (p = 0.012) and more good or fair-grade embryos suitable for implantation (p = 0.013) compared to controls (Tables 5 and 6). These findings suggest antioxidant supplementation improves embryo quality and increases embryo numbers in ICSI. Similar results were reported by Scaruffi et al. (2021), Rago et al. (2017), and Korosi et al. (2017), who found higher embryo quality with various antioxidant-based supplements. Collectively, this evidence supports the role of PROfertil® in enhancing cleavage and embryo quality. In clinical practice, embryos graded as good or fair are prioritized for transfer due to their higher implantation potential and lower miscarriage risk. The improvements in sperm quality observed in phase one of our study may explain the increased number of top quality embryos, contributing to better IVF outcomes.

Blastocyst transfer allows selection of the most viable embryo, which has high implantation potential and

improves outcomes. In our study, blastocyst development was lower in the treatment group, with fewer good-quality blastocysts compared to controls (Table 5). Statistical analysis showed no significant differences between groups, except for the rate of good blastocyst formation. Previous studies by Glujovsky et al. (2016) and Scaruffi et al. (2021) reported higher live birth rates in fresh blastocyst transfer compared to cleavage-stage transfer, which contrasts with our findings. This difference may be explained by the selection of good embryos at the cleavage stage for transfer, while remaining embryos were frozen for future use. As a result, fewer embryos were available for culture to day 5, and those that continued often had lower quality. Embryos with minimal fragmentation are more likely to reach the blastocyst stage, while those with high fragmentation usually arrest after day 3 (Tan et al., 2019). This indicates that research on the effect of oral supplements at the blastocyst stage remains limited, particularly in Malaysia. Further studies are needed to evaluate the efficacy of PROfertil® on blastocyst outcomes.

Additionally, our study revealed fewer poor-quality embryos in the treatment group compared to the control group on both Day 2 and Day 3 (Table 6). A significant difference was observed in the proportion of good- and poor-quality embryos on Day 3 between the treatment and control groups ($p = 0.013$). This finding is important, as embryos graded as 1 (excellent) or 2 (good) have been shown to exhibit higher implantation success rates than poor-quality embryos (Kirillova *et al.*, 2020). The transfer of poor-quality embryos has also been associated with an increased risk of miscarriage and pregnancy loss (Kirillova *et al.*, 2020; Zhu *et al.*, 2014). These results suggest that the improvement in sperm quality achieved after three months of PROfertil® antioxidant treatment may contribute to enhanced early embryonic development, including better fertilization rates, improved cleavage on Day 2, more favorable cleavage on Day 3, and a higher proportion of good-quality embryos on Day 3. Nevertheless, this study has several limitations, including variations in stimulation dosage among female patients and differences in sperm and oocyte quality.

CONCLUSION

Our study demonstrated that three months of PROfertil® supplementation led to measurable improvements in sperm quality, which translated into enhanced early embryonic development. The treatment group exhibited higher fertilization rates, increased day 2 cleavage, enhanced cleavage on day 3, and a greater number of good- to fair-quality embryos compared to controls. These findings highlight the role of antioxidant supplementation in reducing poor-quality embryos and supporting the development of embryos with higher implantation potential. Although improvements in blastocyst formation were limited, the overall outcomes underscore the potential clinical benefits of PROfertil® in assisted reproductive technology. Further studies with larger cohorts and extended follow-up are warranted to confirm its efficacy and impact on live birth outcomes.

ACKNOWLEDGEMENT

We would like to express appreciation to statistician, Dr. Tan Kim Lay, of the Sector for Biostatistics and Data Repository, National Institutes of Health Malaysia (NIH) and Dr. Chua Ker Yang, Paediatrician at Hospital Tunku Azizah for their guidance and contribution to data analysis. My sincere thanks also go to all staff from Kulliyyah of Medicine, Kulliyyah of Allied Health Sciences, Reproductive Medicine Unit Hospital Tunku Azizah, Clinical Research Centre Hospital Tunku Azizah and National Institutes of Health for the support, guidance and access to the laboratory and research facilities.

REFERENCES

Attallah, D., El-Nashar, I. H., Mahmoud, R., Shaaban, O. M., & Salman, S. A. (2013). N-acytelcysteine prior to intrauterine insemination in couples with isolated athenozospermia: a randomized controlled trial. *Fertility and Sterility*, 100(3), S462. http://dx.doi.org/10.1016/j.fertnstert.2013.07.413

Balaban, B., Brison, D., Calderón, G., Catt, J., Conaghan, J., Cowan, L., et al. (2011). Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Hum Reprod.,* 26(6), 1270-1283. https://doi.org/10.1093/humrep/der037

Balakier, H., Sojecki, A., Motamedi, G., & Librach, C. (2016). Impact of multinucleated blastomeres on embryo developmental competence, morphokinetics, and aneuploidy. *Fertility and Sterility,* 106(3), 608-614.e2. https://doi.org/10.1016/j.fertnstert.2016.04.041

Barekat, F., Tavalaee, M., Deemeh, M. R., Bahreinian, M., Azadi, L., Abbasi, H., Rozbahani, S., & Nasr-Esfahani, M. H. (2016). A Preliminary Study: N-acetyl-L-cysteine Improves Semen Quality following Varicocelectomy. *International journal of fertility & sterility*, 10(1), 120–126. https://doi.org/10.22074/ijfs.2016.4777

Bartolacci, A., Pagliardini, L., Makieva, S., Salonia, A.,

Papaleo, E. & Viganò, P. (2018). Abnormal sperm concentration and motility as well as advanced paternal age compromise early embryonic development but not pregnancy outcomes: a retrospective study of 1266 ICSI cycles. *J Assist Reprod Genet*, 35(10), 1897-1903. https://doi.org/10.1007/s10815-018-1256-8

Busetto, G. M., Agarwal, A., Virmani, A., Antonini, G., Ragonesi, G., Del Giudice, F., Micic, S., Gentile, V., & De Berardinis, E. (2018). Effect of metabolic and antioxidant supplementation on sperm parameters in oligo-astheno-teratozoospermia, with and without varicocele: A double-blind placebo-controlled study. *Andrologia*, 50(6), e12927. https://doi.org/10.1111/and.12927

Desai, N., Ploskonka, S., Goodman, L., Attaran, M., Goldberg, J. M., Austin, C., et al. (2016). Delayed blastulation, multinucleation, and expansion grade are independently associated with live-birth rates in frozen blastocyst transfer cycles. *Fertility and Sterility,* 106(6), 1370–1378. https://doi.org/10.1016/j.fertnstert.2016.07.1095

Desch, L., Bruno, C., Luu, M., Barberet, J., Choux, C., Lamotte, M., et al. (2017). Embryo multinucleation at the two-cell stage is an independent predictor of intracytoplasmic sperm injection outcomes. *Fertility and Sterility,* 107(1), 97-103.e4. https://doi.org/10.1016/j.fertnstert.2016.09.022

Dutta, S., Majzoub, A. & Agarwal, A. (2019). Oxidative stress and sperm function: A systematic review on evaluation and management. *Arab J Urol.,* 17(2), 87-97. https://doi.org/10.1080/2090598x.2019.1599624

Egashira, A., Yamauchi, N., Tanaka, K., Mine, C., Otsubo, H., Murakami, M., et al. (2015). Developmental capacity and implantation potential of the embryos with multinucleated blastomeres. *The Journal of Reproduction and Development,* 61(6), 595–600. https://doi.org/10.1262/jrd.2015-052

Imhof, M., Lackner, J., Lipovac, M., Chedraui, P., & Riedl, C. (2012). Improvement of sperm quality after micronutrient supplementation. *E-SPEN Journal*, 7(1), e50–e53. https://doi.org/10.1016/j.eclnm.2011.11.002

Gambera, L., Stendardi, A., Ghelardi, C., Fineschi, B., & Aini, R. (2019). Effects of antioxidant treatment on seminal parameters in patients undergoing in vitro fertilization. *Archivio Italiano Di Urologia e Andrologia*, 91(3), 187–190. https://doi.org/10.4081/aiua.2019.3.187

Gharagozloo, P., Gutierrez-Adán, A., Champroux, A., Noblanc, A., Kocer, A., Calle, A., et al. (2016). A novel antioxidant formulation designed to treat male infertility associated with oxidative stress: Promising preclinical evidence from animal models. *Human Reproduction*, 31(2), 252–262. https://doi.org/10.1093/humrep/dev302

Glujovsky, D., Farquhar, C., Quinteiro Retamar, A. M., Alvarez Sedo, C. R., & Blake, D. (2016). Cleavage stage versus blastocyst stage embryo transfer in assisted reproductive technology. *Cochrane Database of Systematic Reviews,* 5(5), CD002118. https://doi.org/10.1002/14651858.cd002118.pub6

Greco, E., Romano, S., Iacobelli, M., Ferrero, S., Baroni, E., Minasi, M. G., et al. (2005). ICSI in cases of sperm DNA damage: Beneficial effect of oral antioxidant treatment. *Human Reproduction*, 20(9), 2590–2594. https://doi.org/10.1093/humrep/dei091

Kirillova, A., Lysenkov, S., Farmakovskaya, M., Kiseleva, Y., Martazanova, B., Mishieva, N., Abubakirov, A., & Sukhikh, G. (2020). Should we transfer poor quality embryos? *Fertility Research and Practice,* 6(1), 1–7.

Korosi, T., Barta, C., Rokob, K., & Torok, T. (2017). Physiological Intra-Cytoplasmic Sperm Injection (PICSI) outcomes after oral pretreatment and semen incubation with myo-inositol in oligoasthenoteratozoospermic men: results from a prospective, randomized controlled trial. *European Review for Medical and Pharmacological Sciences,* 21(2), 66–72.

Lipovac, M., Nairz, V., Aschauer, J., & Riedl, C. (2021). The effect of micronutrient supplementation on spermatozoa DNA integrity in subfertile men and subsequent pregnancy rate. *Gynecol Endocrinol*, 37(8), 711-715. https://doi.org/10.1080/09513590.2021.1923688

Lipovac, M., Bodner, F., Imhof, M., & Chedraui, P. (2016). Comparison of the effect of a combination of eight micronutrients versus a standard mono preparation on sperm parameters. *Reproductive Biology and Endocrinology*, *14*(1), 1–6. https://doi.org/10.1186/s12958-016-0219-0

Lipovac, M., Bodner, F., Schütz, A., Kurz, H., Riedl, C., Mair, J., et al. (2014). Increased Hyaluronan Acid Binding

Ability of Spermatozoa Indicating a Better Maturity, Morphology, and Higher Dna Integrity After Micronutrient Supplementation. *Emj European Medical Journal,* 60–65.

Majzoub, A., & Agarwal, A. (2018). Systematic review of antioxidant types and doses in male infertility: Benefits on semen parameters, advanced sperm function, assisted reproduction and live-birth rate. *Arab J Urol.,* 16(1), 113-124. https://doi.org/10.1016/j.aju.2017.11.013

Pantos, K., Sfakianoudis, K., Maziotis, E., Rapani, A., Karantzali, E., Gounari-Papaioannou, A., et al. (2021). Abnormal fertilization in ICSI and its association with abnormal semen parameters: A retrospective observational study on 1855 cases. *Asian Journal of Andrology*, 23. https://doi.org/10.4103/aja.aja_84_20

Pino, V., Sanz, A., Valdés, N., Crosby, J., & Mackenna, A. (2020). The effects of aging on semen parameters and sperm DNA fragmentation. JBRA Assisted Reproduction, 24(1), 82–86. https://doi.org/10.5935/1518-0557.20190058

Rago, R., Gallo, M., Dal Lago, A., Licata, E., Paciotti, G., Amodei, M., et al. (2017). Controlled, prospective, observational study on the efficiency and tolerability of a combination of potential Nrf2-inducing antioxidants and micronutrients as pre-treatment for ICSI in dyspermic patients with previous failure. *European Review for Medical and Pharmacological Sciences*, 21(7), 1645–1652.

Rashki Ghaleno, L., Rezazadeh Valojerdi, M., Chehrazi, M., Sahraneshin Samani, F., & Salman Yazdi, R. (2016). Hyaluronic Acid Binding Assay is Highly Sensitive to Select Human Spermatozoa with Good Progressive Motility, Morphology, and Nuclear Maturity. *Gynecologic and Obstetric Investigation*, 81(3), 244–250. https://doi.org/10.1159/000439530

Scaruffi, P., Licata, E., Maccarini, E., Massarotti, C., Bovis, F., Sozzi, F., et al. (2021). Oral antioxidant treatment of men significantly improves the reproductive outcome of ivf cycles. *Journal of Clinical Medicine*, 10(15), 3254 https://doi.org/10.3390/jcm10153254

Schauer (2011). Micronutrients as an alternative to fertility treatment in men with subclinical varicocele. https://profertil.eu/studies/subclinical-varikocele-study-2010/

Smits, R. M., Mackenzie-Proctor, R., Yazdani, A., Stankiewicz, M. T., Jordan, V., & Showell, M. G. (2019). Antioxidants for male subfertility. *The Cochrane database of systematic reviews,* 3(3), CD007411. https://doi.org/10.1002/14651858.cd007411.pub4

Tan, J. H., Chen, J. J., Lim, L. J., & Wong, P. S. (2019). The impact of in vitro human embryo fragmentation on blastocyst development and ploidy using Next-Generation Sequencing (NGS). *Reproductive BioMedicine Online,* 38, e23. https://doi.org/10.1016/j.rbmo.2019.03.039

Tremellen, K., Woodman, R., Hill, A., Shehadeh, H., Lane, M., & Zander-fox, D. (2021). Use of a male antioxidant nutraceutical is associated with superior live birth rates during IVF treatment. *Asian J Androl,* 23(1), 16-23. https://doi.org/10.4103/aja.aja_41_20

Tremellen, K., Miari, G., Froiland, D., & Thompson, J. (2007). A randomised control trial examining the effect of an antioxidant (Menevit) on pregnancy outcome during IVF-ICSI treatment. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 47(3), 216–221. https://doi.org/10.1111/j.1479-828x.2007.00723.x

Vogiatzi, P., Pouliakis, A., Sakellariou, M., Athanasiou, A., Colaghis, A., Finelli, R., et al. (2022). Male Age and Progressive Sperm Motility Are Critical Factors Affecting Embryological and Clinical Outcomes in Oocyte Donor ICSI Cycles. *Reprod Sci.*, 29(3), 883-895. https://doi.org/10.1007/s43032-021-00801-1

Zhu DL, Zhang HG, Wang RX, Jiang YT, & Liu RZ (2019). Re-evaluation of the value of sperm morphology in classica l in vitro fertilization in a Northeastern Chinese population. *Journal of International Medical Research,* 47(9), 4134–4142. https://doi.org/10.1177/0300060519860324

Zhu J, Lian Y, Li M, Chen L, Liu P, & Qiao J. (2014). Does IVF cleavage stage embryo quality affect pregnancy complications and neonatal outcomes in singleton gestations after double embryo transfers? *Journal of Assisted Reproduction and Genetics*, 31(12), 1635–1641. https://doi.org/10.1007/s10815-014-0351-8

ORIGINAL ARTICLE

Artificial Intelligence in Oncology for Early Detection and Intervention: Clinical and Operational Insights from Malaysia

Nik Muhammad Alif Bin Mohd Azmi[1] and Mardhiah Mohammad[1,*]

[1]Department of Biomedical Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia

Keywords:
artificial intelligence; Malaysia; oncology; applications; outcomes

ABSTRACT

Background: Cancer is a significant public health challenge in Malaysia and become a major contributor to morbidity and mortality. Early detection and timely intervention are essential to improving survival rates, yet conventional oncology services often face delays and resource constraints. Artificial Intelligence (AI) has emerged as a transformative tool capable of enhancing diagnostic accuracy, treatment planning, and patient monitoring. **Methods:** A scoping review was conducted using the PRISMA-ScR framework. Articles published between 2015 and 2025 were retrieved from PubMed, Google Scholar, and IEEE Xplore. Studies were screened and selected based on predefined inclusion and exclusion criteria, before being analysed into three clinical domains: diagnosis and screening, treatment planning, and monitoring and prognosis. **Results:** Fifteen studies were included. AI applications demonstrated improved diagnostic sensitivity and specificity, radiotherapy planning, and accurate survival prediction models. Operationally, AI contributed to enhanced workflow efficiency, cost reductions, and better decision-making support, particularly in resource-limited settings. **Conclusion:** AI shows significant promise for advancing early detection and intervention in Malaysian oncology, delivering both clinical and operational benefits.

INTRODUCTION

Cancer continues to be one of the leading causes of morbidity and mortality in Malaysia. According to the Malaysian National Cancer Registry, breast, colorectal, and lung cancers consistently rank among the most prevalent malignancies in the country (Haron et al., 2022; Sharma et al., 2024). The increasing cancer burden, driven by population aging, urbanization, and changing dietary and lifestyle habits, has placed immense strain on the nation's healthcare system. Delivering timely and accurate cancer diagnosis, personalized treatment, and long-term monitoring remains a challenge, particularly in rural or resource-limited settings where access to specialized oncology care is restricted (Lim et al., 2021; Kadir et al., 2022; Hassan et al., 2024).

To address these challenges, AI has emerged globally as a transformative force in oncology. AI refers to computational algorithms and systems designed to mimic human cognitive functions such as learning, pattern recognition, and decision-making (Mahusin et al., 2024). In the context of cancer care, AI technologies, especially those based on machine learning (ML) and deep learning (DL) are being deployed to support clinical decision-making, automate repetitive tasks, and improve data interpretation (Phang et al., 2024). By processing large and complex datasets with high precision, AI can facilitate early detection, personalise treatment strategies, and support continuous monitoring, ultimately strengthening allied health contributions to cancer care.

In Malaysia, AI integration into oncology is still in its early stages. A limited number of academic and clinical studies have begun to explore the feasibility of using AI for cancer care, often through pilot projects and prototype models (Sachithanandan et. al, 2024). Most of these efforts focus on breast and colorectal cancers, reflecting national cancer trends. Institutions such as University Malaya Medical Centre (UMMC), University Tunku Abdul Rahman Malaysia (UTAR) and University Tun Hussein Onn Malaysia have been instrumental in developing AI-based models using local datasets (Ganggayah et. al, 2021; Lim et al., 2021; Moksem & Newaz, 2024). However, these projects often remain within the scope of research and have not yet translated into widespread clinical practice.

Despite these limitations, the Malaysian government's push toward digital transformation under the National Artificial Intelligence Roadmap 2021–2025 provides a supportive backdrop for future AI implementation in healthcare (Phang et al., 2024). There is also growing interest among researchers and policymakers in harnessing AI to improve healthcare equity, particularly in

* *Corresponding author.*
E-mail address: mmoh@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

underserved areas. Given the growing cases of cancer in Malaysia, AI offers a powerful means to enhance diagnostic accuracy, personalize treatment strategies, and streamline patient monitoring.

This scoping review aims to synthesise existing literatures on the clinical and operational impact of AI applications in oncology within Malaysian healthcare context, with a focus on early detection and intervention. It is hoped that the insights will guide researchers, clinicians, and policymakers in understanding the readiness, addressing the challenges, and leveraging the opportunities for responsible AI integration in Malaysia's cancer care ecosystem.

MATERIAL AND METHOD

Study Design

A scoping review was conducted following to the five-stage framework by Arksey & O'Malley (2005), incorporating the checklist from the Preferred Reporting Items for Systematic Reviews and Meta- Analyses extension for scoping reviews (PRISMA-Scr) (Tricco et al., 2018).

Search Strategy

Relevant research articles were selected and retrieved from credible online databases such as Google Scholar, PubMed, and IEEE Xplore. The search was guided by specific keywords, including 'Artificial Intelligence', 'Malaysia', 'Oncology', 'Applications', 'Outcomes', 'Clinical Outcomes', and 'Operational Outcomes'. Boolean operators (AND, OR) were used to combine these keywords appropriately. The 'AND' operator was used to link distinct keywords into comprehensive search strings, while the 'OR' operator was used to group the synonyms of each keyword. Examples of keywords combined using 'OR' include: "Artificial Intelligence, Machine Learning, Neural Networks, Deep Learning"; "Oncology, Cancer Study"; "Healthcare, Hospital, Institute, Clinic"; "Application, Utilisation, Usage, Integration"; and "Outcomes, Results, Impacts".

Inclusion and Exclusion Criteria

The review included study articles of any research designs which is quantitative (e.g., experimental, cross-sectional), qualitative (e.g., interviews, case studies), mixed-methods, and secondary research such as literature and systematic reviews. This broad criterion allowed diverse evidence, including technical validation studies and retrospective analyses relevant to AI applications in oncology. Eligble studies were published between 2015 and 2025, written in English, matched to the research keywords that available in full-text. Exclusions included non-English articles published before 2015 and incomplete works such as abstracts, conference papers, or posters.

Screening and Selection

Titles and abstract were screened before full-text assessment were conducted to map the findings on AI applications and outcomes in oncology. Discrepancies were resolved through discussion among the reviewers.

Data Extraction and Synthesis

The data for AI usages and outcomes were classified based on key areas of diagnostic and screening, treatment planning, and monitoring and prognosis. The findings were further synthesised, which the clinical and operational outcomes that have been reported from the application of AI in oncology in Malaysia across these clinical domains were determined.

RESULTS

A total of 1586 articles were identified for this review via three electronic databases. Following the removal of 95 duplicates, 1491 articles were selected for the screening process. The screening phase resulting in the exclusion of 1412 articles were excluded as it is not related with research keywords in the title, abstract, and publication year. Next, the articles were assessed for eligibility based on inclusion criteria and exclusion criteria which 159 articles were excluded as it does not have full text access, also excluded articles that only have abstract, conference, or incomplete papers. Overall, 15 studies were identified as being eligible for full text evaluation (Figure 1).

Study characteristics

These articles highlighted that AI applications in Malaysian oncology primarily focus on three clinical domains: diagnosis and screening, treatment planning, and m onitoring and prognosis. Across domains, reported advantages included faster diagnosis, reduced costs, and improved workflow efficiency, although challenges such as small datasets, limited local validation, and integration barriers persisted. Together, these findings provide a comprehensive view of how AI is shaping cancer care in Malaysia.

Study type

Of the 15 studies included, 6 were retrospective in design. Only two study employed a prospective study design. For three studies, the study design could not be clearly determined as either retrospective or prospective. Among the included studies, five were conducted within hospital and cancer centre settings, four were undertaken in university-based research environments, and one study

was carried out in a Positron Emission Tomography–Computed Tomography scan **(**PET-CT) scan facility. Additionally, seven studies drew upon global datasets, with several explicitly incorporating data contributions from Malaysia, thereby enhancing the relevance of findings to the local context. This distribution of study settings and data sources highlights both the diversity and representativeness of the available evidence, while also underscoring potential variations in generalisability and applicability across different healthcare contexts (Supplement Table 1).

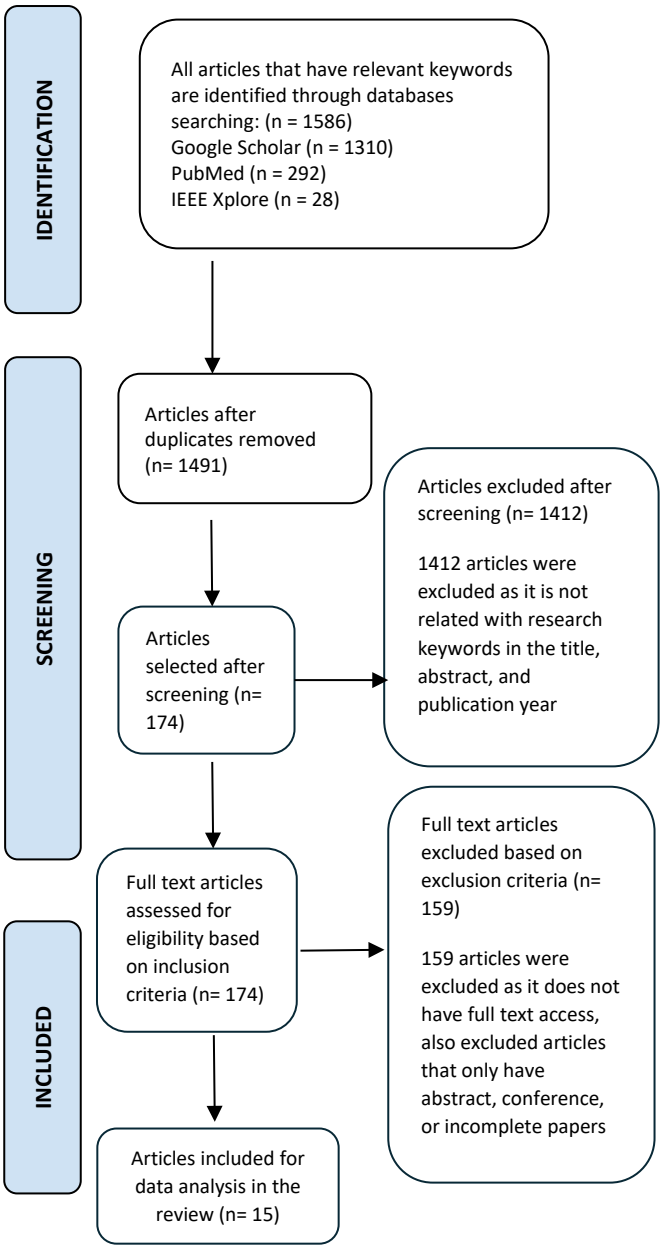


Figure 1. PRISMA-ScR flowchart

Diagnosing and Screening

AI applications in oncology have demonstrated remarkable improvements in diagnostic accuracy across multiple cancer types in Malaysia. For breast cancer, Lim et al. (2021) reported that a hybrid Support Vector Machine-Articial Neural Network (SVM-ANN) model achieved 99.4% diagnostic accuracy, with sensitivity of 100% and specificity of 97.7%. Similarly, Ling et al. (2022) showed that integrating circulating micro RNA (miRNA) profiles with machine learning achieved over 90% accuracy in breast cancer detection, while Zhang et al. (2024) demonstrated that big-data AI models incorporating multi-omics datasets could detect breast tumors with 99% accuracy. In lung cancer, AI-assisted chest X-rays improved sensitivity from 66.4% to 74.7% and reduced false-positive rates from 0.25 to 0.18 (Sachithanandan et al., 2024). For gastric cancer, hybrid deep learning models reached 99.49% classification accuracy (Mirza et al., 2023). Collectively, these findings highlight that AI enhances early detection and diagnostic precision, which is critical for timely intervention.

Beyond clinical performance, AI has also significantly improved operational efficiency in diagnostic workflows. Lim et al. (2021) found that hybrid SVM-ANN models reduced the manual burden on radiologists and accelerated diagnostic turnaround times. Ling et al. (2022) reported that combining circulating miRNA with AI reduced reliance on invasive tissue biopsies and shortened the diagnostic process. In lung cancer, AI-assisted chest X-rays facilitated more efficient referrals for low-dose CT scans and lowered unnecessary follow-ups (Sachithanandan et al., 2024). Similarly, for gastric cancer, Mirza et al. (2023) suggested that AI could potentially reduce unnecessary endoscopies, optimizing resource utilization. These operational benefits demonstrate that AI not only improves accuracy but also enhances workflow efficiency, supporting timely patient management and intervention.

Treatment Planning

AI has significantly improved the precision and effectiveness of oncology treatment planning in Malaysia. In radiotherapy, Moksem and Newaz (2024) reported that AI-based planning reduced tumor contouring errors by 60%, while Aggarwal et al. (2023) found that 95% of AI-generated radiotherapy plans met international clinical standards. For chemotherapy, AI-optimized regimens led to a 20% reduction in treatment-related toxicity (Abdul Rasool Hassan et al., 2025). AI has also enhanced targeted therapies, with Poh et al. (2023) demonstrating that AI-supported selection of Anaplastic Lymphoma Kinase (ALK) -targeted therapy increased patient survival to 62 months with first-generation ALK inhibitors. Furthermore, Faiz et al. (2024) showed that AI-assisted genetic profiling enabled more precise targeted therapies for gastric cancer patients. Collectively, these findings indicate that AI not only improves treatment accuracy but also optimizes

clinical outcomes across radiotherapy, chemotherapy, and targeted therapy.

Beyond clinical performance, AI has also significantly improved operational efficiency in diagnostic workflows. Lim et al. (2021) found that hybrid SVM-ANN models reduced the manual burden on radiologists and accelerated diagnostic turnaround times. Ling et al. (2022) reported that combining circulating miRNA with AI reduced reliance on invasive tissue biopsies and shortened the diagnostic process. In lung cancer, AI-assisted chest X-rays facilitated more efficient referrals for low-dose CT scans and lowered unnecessary follow-ups (Sachithanandan et al., 2024). Similarly, for gastric cancer, Mirza et al. (2023) suggested that AI could potentially reduce unnecessary endoscopies, optimizing resource utilization. These operational benefits demonstrate that AI not only improves accuracy but also enhances workflow efficiency, supporting timely patient management and intervention.

Monitoring and Prognosis

AI applications have markedly enhanced monitoring and prognostic accuracy in oncology. For breast cancer, Kadir et al. (2022) demonstrated that an ANN model achieved 84.5% accuracy in predicting 5-year survival, while Ganggayah et al. (2021) reported that a Random Forest model reached 92.5% survival prediction accuracy. AI-based volumetric tumor monitoring also improved therapy response assessment, as shown by Mahdi et al. (2024), providing more precise and timely evaluations of tumor progression. In gastric cancer, AI models outperformed conventional Model for End-Stage Liver Disease (MELD) scores in predicting mortality (Faiz et al., 2024), and in lung cancer, AI-assisted follow-up enabled earlier detection of malignant nodules (Sachithanandan et al., 2024). Collectively, these studies indicate that AI supports more accurate prognostic assessments, facilitating earlier interventions and tailored patient management.

AI has also improved operational efficiency in patient monitoring and prognosis. Personalized follow-up plans for breast cancer patients were facilitated by AI predictions, optimizing resource allocation (Kadir et al., 2022), while longitudinal follow-up of over 1,000 patients was streamlined using Random Forest (RF) survival models (Ganggayah et al., 2021). Volumetric tumor monitoring reduced manual measurement bias, enhancing workflow efficiency in therapy assessments (Mahdi et al., 2024). AI-assisted prognostic models also enabled better ICU triaging for gastric cancer patients (Faiz et al., 2024) and decreased missed diagnoses during lung nodule follow-up (Sachithanandan et al., 2024). These operational improvements highlight AI's capacity to optimize patient monitoring, support timely clinical decisions, and enhance overall healthcare efficiency.

DISCUSSION

To our knowledge, this is the most recent scoping review examining the applications of artificial intelligence in oncology within the Malaysian context. A total of 15 studies were identified, covering diverse domains including diagnosis and screening, treatment planning, and monitoring and prognosis. Among these, diagnostic applications were the most frequently reported, reflecting the growing use of AI to support cancer detection and early intervention.

Majority of the studies relied on retrospective data sources. Only one prospective case-control design was reported, and no prospective cohort studies or randomized controlled trials were identified. This highlights the methodological gap in robust validation of AI models in real-world clinical practice. The findings underscore the need for future research to move beyond retrospective analyses and adopt prospective, multi-center, or randomized designs. Such approaches are essential to strengthen evidence on the reliability, generalizability, and clinical utility of AI applications for cancer care in Malaysia.

Nevertheless, AI has played a transformative role across all domains of oncology care, including diagnosis and screening, treatment planning, and monitoring and prognosis; enabling more precise, personalised, and efficient patient management. In Malaysia, AI applications have demonstrated substantial improvements in early detection, optimising therapy selection and radiation dosing, personalising chemotherapy regimens, and predicting patient outcomes, with benefits reported for both clinical and operational performance.

The majority of the studies concentrated on the clinical domain of diagnosis and screening, with fewer studies addressing treatment planning and monitoring and prognosis. Most studies relied on retrospective datasets, and only a few used locally sourced Malaysian clinical data, constraining the generalizability of findings to the national healthcare landscape (Mirza et al., 2024).

Some articles explored AI-driven optimization of chemotherapy regimens; however, these applications remain largely in academic or pilot stages, without large-scale deployment in clinical environments. Whereas, in the monitoring and prognosis domain, AI models were developed to predict patient survival rates, disease

progression, and recurrence risks, often with high sensitivity and specificity. Nevertheless, the lack of prospective studies limits confidence in their real-world reliability (Kadir et al., 2022; Aggarwal et al., 2023).

Most AI applications were implemented in academic or simulation-based evaluations rather than routine clinical practice. For instance, Moksem and Newaz (2024) carried out a clinical system evaluation study which applying AI-driven radiotherapy planning into radiation oncology departments in UNIRAZAK and UMMC. Similarly, a study conducted by Abdul Rasool Hassan et al. (2025) shown a predictive AI model that has capability in predicting the chemotherapy toxicity by relying on the patient records.

In lung cancer, AI research has focused on improving screening accuracy, radiotherapy precision, and treatment decision support. AI-enhanced imaging systems have demonstrated significant promise in early detection. For example, Sachithanandan et al. (2024) showed that AI-assisted interpretation of chest X-rays combined with low-dose computed tomography (LDCT) increased the sensitivity for detecting early-stage non–small cell lung cancer (NSCLC) and reduced false positives, thereby improving population-based screening efficiency.

Whereas for treatment planning, AI has revolutionized radiotherapy by improving dose precision and minimizing radiation exposure to surrounding tissues. Mahdi et al. (2024) introduced an AI-enabled fusion transformer model for PET/CT-based segmentation in head and neck and lung tumors, achieving superior performance compared to manual delineation. Similarly, Abdul Rasool Hassan et al. (2025) reported that AI-assisted chemotherapy development and selection tools enhanced therapeutic effectiveness while reducing toxicity profiles.

Furthermore, a retrospective study conducted by Poh et al. (2023) in 18 hospitals in Malaysia which a decision support system for ALK positive non-small cell lung cancer (NSCLC) using a molecular diagnostic was combined with AI algorithms for drug matching. The model was tested within clinical oncology services and it has presented evidence that AI-supported decision-making in ALK-targeted therapy selection extended median survival to 62 months in patients receiving first-generation ALK inhibitors. This underscores the potential of AI to directly influence long-term patient outcomes when integrated into precision oncology workflows.

Various AI model have also been applied to assist in early detection, treatment optimization, and outcome prediction of breast cancer. Diagnostic models using hybrid SVM–ANN model (Lim et al., 2021) and multi-omics big data approaches (Zhang et al., 2024) have achieved diagnostic accuracies exceeding 95%, demonstrating AI's capability to identify malignant lesions from mammography and histopathology images with minimal human bias. These innovations support radiologists in large-scale screening programs by reducing false negatives and enhancing the detection of early-stage breast lesions.

AI also contributes significantly to treatment planning and prognostic modeling by developing machine learning models that predict survival probabilities and guide treatment selection based on clinical and pathological parameters (Kadir et al., 2022). Such model enable oncologists to tailor therapeutic intensity to patient-specific profiles and to identify high-risk individuals who may benefit from more intensive follow-up.

Meanwhile, in colorectal cancer, AI adoption in colorectal cancer has been applied in endoscopic diagnostics and prognostic prediction. AI-enabled colonoscopy systems using deep convolutional neural networks have shown remarkable efficacy in real-time polyp and adenoma detection, achieving sensitivities exceeding 96% (Faiz et al., 2024). These systems improve lesion recognition and procedural safety, contributing to earlier diagnosis and reduced colorectal cancer mortality.

Several systemic barriers limit the broader implementation of AI in Malaysian healthcare. Among the most significant are the lack of standardized and high-quality local datasets required for training AI models, as well as inadequate digital infrastructure and high-performance computing capabilities in public hospitals (Poh et al., 2023; Moksem & Newaz, 2024). Additionally, many clinicians remain unfamiliar with AI technology and lack the training to effectively interpret or validate AI-generated recommendations (Aggarwal et al., 2023). On a regulatory level, Malaysia does not yet have dedicated policies or legal frameworks that govern the clinical use of AI, raising ethical concerns about patient data privacy, algorithmic transparency, and liability in the event of clinical error (Putera et al., 2021; Phang et al., 2024).

In another study which Aggarwal et al. (2023) reported that 95% of AI-generated radiotherapy plans met international clinical standards, achieving reductions in planning time from weeks to minutes, which is crucial for timely cancer treatment initiation. Though it has shown a good result in the radiotherapy, this system however was limited to facilities with advanced infrastructure and personnel trained in AI systems.

In contrast with other country, for example, Shi et al. (2025) provided a comprehensive review of AI applications in high-dose-rate (HDR) brachytherapy for cervical cancer, highlighted several pilot and retrospective studies showing promising results in reducing planning time, increasing plan conformity, and improving dosimetric outcomes, but also emphasized the lack of standardization, limited datasets, and high technical barriers for clinical integration. So, while Aggarwal et al. (2023) presented real-world evidence of AI integration in external beam radiotherapy (RT) planning at scale, demonstrating operational success and feasibility, Shi et al. (2025) offered a technical roadmap for AI applications in High-Dose-Rate (HDR) brachytherapy, underlining theoretical advantages but with limited clinical adoption. Both has affirmed AI's potential to revolutionise radiotherapy planning, but differ in maturity as Aggarwal's model is validated in clinical settings, while Shi's represents a still-developing frontier.

AI is transforming cancer care by improving survival prediction, follow-up efficiency, and clinical decision-making**,** offering new avenues for personalized treatment and patient management. Comparing Malaysia and China provides valuable insight into how different healthcare settings, population needs, and clinical priorities shape the design and application of AI-based prognostic models**.**

In Malaysia, the iSurvive model developed by Ganggayah et al. (2021) represents a holistic, long-term care approach for survivors**,** integrating demographic, lifestyle, and clinical variables to predict survival outcomes and support ongoing patient engagement. This aligns with Malaysia's growing emphasis on survivorship programs and long-term quality-of-life monitoring for breast cancer patients. In contrast, the model introduced by Zhen et al. (2024) in China adopts a focused, short-term mortality risk strategy for targeted intervention, emphasizing modifiable lifestyle factors to identify high-risk patients for early management. This reflects China's healthcare priority of early mortality reduction in a population with diverse lifestyle-related risk factors.

The comparative analysis highlights how AI's flexibility allows adaptation to different national priorities, from survivorship monitoring in Malaysia to mortality prediction in China. Despite their strong predictive performance, iSurvive achieving over 90% survival prediction accuracy and Zhen's model demonstrating high AUC values for 1- and 3-year mortality. Although both models demonstrate strong predictive performance, further external validation, improved data completeness, and seamless software integration are essential before widespread adoption. These findings highlight AI's potential as a flexible and powerful tool in oncology, capable of supporting diverse clinical goals across different healthcare systems.

CONCLUSION

This scoping review explored the applications of Artificial Intelligence (AI) in oncology within the Malaysian context, categorising evidence into three clinical domains—diagnosis and screening, treatment planning, and monitoring and prognosis—and assessing their clinical and operational outcomes.

The findings demonstrate that AI has delivered measurable benefits in enhancing diagnostic accuracy, optimising treatment workflows, and enabling personalised patient monitoring. In diagnosis and screening, AI models such as convolutional neural networks (CNNs) and support vector machines (SVMs) have improved sensitivity, specificity, and early cancer detection rates, particularly for breast, colorectal, and lung cancers. In treatment planning, AI has reduced radiotherapy planning times, enhanced dose prediction accuracy, and supported personalised chemotherapy selection. In monitoring and prognosis, AI-driven survival prediction models, recurrence risk tools, and mobile health platforms have facilitated continuous follow-up and proactive patient management.

From an operational perspective, AI has contributed to improved workflow efficiency, reduced specialist workloads, prioritised urgent cases, and lowered healthcare costs in resource-constrained settings. These benefits are especially relevant for Malaysia, where oncology services often face constraints in manpower, infrastructure, and access to specialised care.

However, the review also identified several barriers limiting widespread AI adoption. Infrastructure gaps, particularly in high-performance computing and digital health integration, remain significant. The scarcity of large, standardised, and representative Malaysian datasets restricts the robustness and generalisability of AI models. Clinician readiness and training in AI use are limited, while ethical concerns, such as patient data privacy, algorithm transparency, and liability in medical decision-making require urgent attention.

The limitations of this review include its restriction to studies published between 2015 and 2025, which may have excluded earlier relevant works. Only English-language publications with full-text access were included, potentially omitting valuable data from non-English sources or conference proceedings. Additionally, the review focused solely on Malaysian oncology, meaning

that broader global trends and innovations were not assessed in depth. Finally, as a scoping review, the study did not conduct formal quality appraisal of included articles, which may limit the interpretability of the findings.

Nevertheless, AI holds substantial promise in transforming cancer care delivery in Malaysia by enhancing accuracy, efficiency, and personalisation across the cancer care continuum. Realising this potential will require strategic investments in infrastructure, capacity building, data governance, and regulatory frameworks. Collaborative efforts between healthcare providers, researchers, policymakers, and technology developers will be critical to ensuring that AI adoption is ethical, equitable, and sustainable, ultimately improving outcomes for cancer patients nationwide.

AI has demonstrated substantial potential to enhance oncology care in Malaysia across diagnosis and screening, treatment planning, and monitoring/prognosis. By improving early detection, optimizing therapy selection, and enabling precise prognostic assessment, AI supports timely interventions and personalized patient management. Operationally, these applications streamline workflows, reduce clinician burden, and improve resource allocation.

Despite promising results, most studies remain retrospective, pilot-based, or simulation-focused, with limited use of locally sourced Malaysian data. Future research should prioritize prospective trials, local dataset development, and real-world implementation to ensure clinical reliability and scalability.

Integrating AI into oncology care aligns with the allied health mandate to promote prevention, early detection, and efficient intervention, ultimately improving patient outcomes and supporting evidence-based practice in Malaysia's healthcare system.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Abdul Rasool Hassan, B. A. R., Mohammed, A. H., Hallit, S., Malaeb, D., & Hosseini, H. (2025). Exploring the role of artificial intelligence in chemotherapy development, cancer diagnosis, and treatment: Present achievements and future outlook. *Frontiers in Oncology*, *15*, 1475893. https://doi.org/10.3389/fonc.2025.1475893

Aggarwal, A., Court, L. E., Hoskin, P., Jacques, I., Kroiss, M., Laskar, S., Lievens, Y., Mallick, I., Malik, R. A., Miles, E., Mohamad, I., Murphy, C., Nankivell, M., Parkes, J., Parmar, M., Roach, C., Simonds, H., Torode, J., Vanderstraeten, B., & Langley, R. (2023). ARCHERY: a prospective observational study of artificial intelligence-based radiotherapy treatment planning for cervical, head and neck and prostate cancer – study protocol. *BMJ Open, 13*(12), e077253. https://doi.org/10.1136/bmjopen-2023-077253

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal Social Research Methodology, 8* (1), 19–32. https://doi.org/10.1080/1364557032000119616

Faiz, A., Nurul Aisyah, N., Hafiz, M., Siti Zulaikha, S., & Hakimi, I. (2024). *Evaluating the role of artificial intelligence in enhancing diagnostic accuracy and personalized treatment of gastrointestinal and hepatobiliary diseases* [Preprint]. ResearchGate. https://www.researchgate.net/publication/386382603

Ganggayah, M. D., Dhillon, S. K., Islam, T., Kalhor, F., Chiang, T. C., Kalafi, E. Y., & Taib, N. A. (2021). An artificial intelligence-enabled pipeline for medical domain: Malaysian breast cancer survivorship cohort as a case study. *Diagnostics, 11*(8), 1492. https://doi.org/10.3390/diagnostics110814

Haron, M. A., Yusof, S. N., Raman, S., Hamzah, N., Muhamat, S. M., Bahtiar, B., *et al.* (2024). *The Malaysia National Cancer Registry Report*. Putrajaya: Ministry of Health Malaysia

Hassan, J., Saeed, S. M., Deka, L., Uddin, M. J., & Das, D. B. (2024). Applications of machine learning (ML) and mathematical modeling (MM) in healthcare with special focus on cancer prognosis and anticancer therapy: Current status and challenges. *Pharmaceutics, 16*(2), 260. https://doi.org/10.3390/pharmaceutics16020260

Kadir, M. N. N. A., Yaacob, N. M., Yusof, S. N., Hadi, I. S. A., Musa, K. I., Isa, S. A. M., Bahtiar, B., Adam, F., Yahya, M. M., & Hairon, S. M. (2022). Development of predictive models for survival among women with breast cancer in Malaysia. *International Journal of Environmental Research and Public Health, 19*(22), 15335. https://doi.org/10.3390/ijerph192215335

Lim, T. S., Tay, K. G., Huong, A., & Lim, X. Y. (2021). Breast cancer diagnosis system using hybrid support vector

machine-artificial neural network. *International Journal of Power Electronics and Drive Systems/International Journal of Electrical and Computer Engineering, 11*(4), 3059 –3069. https://doi.org/10.11591/ijece.v11i4.pp3059-3069

Ling, L., Aldoghachi, A. F., Chong, Z. X., Ho, W. Y., Yeap, S. K., Chin, R. J., Soo, E. Z. X., Khor, J. F., Yong, Y. L., Ling, J. L., Yan, N. S., & Ong, A. H. K. (2022). Addressing the clinical feasibility of adopting circulating miRNA for breast cancer detection, monitoring and management with artificial intelligence and machine learning platforms. *International Journal of Molecular Sciences, 23*(23), 15382. https://doi.org/10.3390/ijms232315382

Mahdi, M. A., Ahamad, S., Saad, S. A., Dafhalla, A., Qureshi, R., & Alqushaibi, A. (2024). Weighted fusion transformer for dual PET/CT head and neck tumor segmentation. *IEEE Access, 12,* 110905–110919. https://doi.org/10.1109/access.2024.3439439

Mahusin, N., Sallehudin, H., & Satar, N. S. M. (2024). Malaysia public sector challenges of implementation of artificial intelligence (AI). *IEEE Access, 12,*121035–121051. https://doi.org/10.1109/access.2024.3448311

Mirza, O. M., Alsobhi, A., Hasanin, T., Ishak, M. K., Karim, F. K., & Mostafa, S. M. (2023). Computer aided diagnosis for gastrointestinal cancer classification using hybrid rice optimization with deep learning. *IEEE Access, 11,* 76321–76329. https://doi.org/10.1109/access.2023.3297441

Mohd Nor, N. A. M., Taib, N. A., Saad, M., Zaini, H. S., Ahmad, Z., Ahmad, Y., & Dhillon, S. K. (2019). Development of electronic medical records for clinical and research purposes: the breast cancer module using an implementation framework in a middle income country- Malaysia. *BMC Bioinformatics, 19*(Suppl 13), Article 2406. https://doi.org/10.1186/s12859-018-2406-9

Moksem, N., H., & Newaz, F., T. (2024). Impact of artificial intelligence (AI) integration in radiotherapy treatment planning using lean management practice. *International Journal of Business and Technology Management, 6,* (1) 140–154. Retrieved from https://myjms.mohe.gov.my/index.php/ijbtm

Phang, K. C., Ng, T. C., Singh, S. K. G., Voo, T. C., & Alvis, W. A. (2024). Navigating Artificial intelligence in Malaysian healthcare: Research developments, ethical dilemmas, and governance strategies. *Asian Bioethics Review.* https://doi.org/10.1007/s41649-024-00314-4

Poh, M. E., How, S. H., Ho, G. F., Pang, Y. K., Hasbullah, H. H., Tho, L. M., Nor, I. M., Lim, B. C., Ho, K. F., Thiagarajan, M., Samsudin, A., Omar, A., Ong, C. K., Soon, S. Y., Tan, J. Y. K., & Abidin, M. a. Z. (2023). Real-world treatment and outcomes of ALK positive metastatic non–small cell lung cancer in a Southeast Asian country. *Cancer Management and Research, 15,* 31–41. https://doi.org/10.2147/cmar.s393729

Putera, N. S. F. M. S., Abdullah, S. M., Rahman, N. A., Hassan, R. A., Saripan, H., & Haryanto, I. (2021). Malaysian medical device regulation for artificial intelligence in healthcare: Have all the pieces fallen into position? *Environment-Behaviour Proceedings Journal, 6(*16), 137–144. https://doi.org/10.21834/ebpj.v6i16.2635

Sachithanandan, A., Lockman, H., Azman, R. R., Tho, L. M., Ban, E. Z., & Ramon, V. (2024). The potential role of artificial intelligence-assisted chest X-ray imaging in detecting early-stage lung cancer in the community-a proposed algorithm for lung cancer screening in Malaysia. *Medical Journal of Malaysia, 79*(1), 9–14. https://pubmed.ncbi.nlm.nih.gov/38287751

Sharma, R., Abbastabar, H., Abdulah, D. M., Abidi, H., Abolhassani, H., Abrehdari-Tafreshi, Z., Absalan, A., Ali, H. A., Abu-Gharbieh, E., Acuna, J. M., Adib, N., Adnani, Q. E. S., Aghaei, A., Ahmad, A., Ahmad, S., Ahmadi, A., Ahmadi, S., Ahmed, L. A., Ajami, M., Fukumoto, T.(2024). Temporal patterns of cancer burden in Asia, 1990–2019: A systematic examination for the Global Burden of Disease 2019 study. *The Lancet Regional Health - Southeast Asia, 21,* 100333. https://doi.org/10.1016/j.lansea.2023.100333

Shi, J., Chen, J., He, G., & Peng, Q. (2025). Artificial intelligence in high-dose-rate brachytherapy treatment planning for cervical cancer: A review. *Frontiers in Oncology, 15,* 1507592. https://doi.org/10.3389/fonc.2025.1507592

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L, Aldcroft, A., Wilson, M. G., Garritty, C., Lewin, S., Godfrey, C. M., Macdonald, M. T., Langlois, E. V., Soares-Weiser, K., Moriarty, J., Clifford, T., Tunçalp, Ö., & Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine., 169(7)* 467-473. https://doi.org/10.7326/M18-0850

Zhang, H., Hussin, H., Hoh, C., Cheong, S., Lee, W., &

Yahaya, B. H. (2024). Big data in breast cancer: Towards precision treatment. *Digital Health, 10*, Article 20552076241293695. https://doi.org/10.1177/20552076241293695

Zhang, L., Yang, D., Ye, X., & Bai, C. (2025). Successful application of artificial intelligence assisted analysis of invasive pulmonary adenocarcinoma less than 6 mm in size: A case report and literature review. *The Clinical Respiratory Journal, 19*(5). https://doi.org/10.1111/crj.70073

Zhen, M., Chen, H., Lu, Q., Li, H., Yan, H., & Wang, L. (2024). Machine learning-based predictive model for mortality in female breast cancer patients considering lifestyle factors. *Cancer Management and Research, 16,* 1253–1265. https://doi.org/10.2147/CMAR.S460811

Supplementary Table 1

Table 1: Characteristics of articles reviewed: Types of study designs and study findings

Authors	Data Collection Method	Study Setting	AI Outcome in Diagnosis & Screening	AI Outcome in Treatment Planning	AI Outcome in Monitoring & Prognosis	Advantages (Clinical/Operational)	Limitations
Mohd Nor et al. (2019)	Retrospective	University Malaya Medical Centre (UMMC), Malaysia	Standardized data improved diagnosis; time to diagnosis reduced to 16.8 days	Enhanced multidisciplinary treatment planning with data sharing and e-prescription	6,974 follow-ups tracked; survival analysis supported	Faster diagnosis, fewer errors, better data quality	Resistance by clinicians; partial interoperability; limited AI integration
Sachithanandan et al. (2024)	Neither retrospective not prospective	Klang Valley, Malaysia	AI-assisted X-rays improved sensitivity (66.4%→74.7%)	Prioritized high-risk cases for LDCT; reduced unnecessary scans	Tracked nodules for early intervention	Cost-effective, scalable, faster diagnosis	Poor follow-up adherence; technical barriers; clinician skepticism
Ganggayah et. al, 2021	Prospective	UMMC, Malaysia	ML identified survival predictors (BMI, stage, etc.)	AI-guided personalized treatment strategies	Longitudinal QoL and survival prediction	92.5% survival prediction accuracy; streamlined data	Missing data issues; software integration problems; external validation needed

Faiz et. al, 2024	Neither retrospective not prospective	Global (including Malaysia's hospital and research settings)	AI in endoscopy improved colorectal polyp detection (96.4% sensitivity)	AI-guided genetic therapy personalization	AI predicted cirrhosis mortality better than MELD score	Faster endoscopy, reduced costs, optimized ICU resources	Data heterogeneity; black-box AI issues; regulation challenges
Hassan et al. (2024)	retrospective	Global (including Malaysia)	ML predicted oral cancer at 78.95% accuracy	AI-guided drug delivery models	ML improved telemedicine for rural patients	Faster, cheaper drug discovery; high telemedicine satisfaction	Small datasets; generalizability concerns; limited direct clinical trials
Poh et al. (2023)	Retrospective study (MNCR data)	18 Hospitals in Malaysia	Potential AI use suggested for faster ALK detection (not implemented yet)	AI could optimize ALK inhibitor selection	AI could predict survival based on big data	ALK inhibitors extended survival vs. chemotherapy	Limited sample; cost barriers; delayed genetic testing
Kadir et al. (2022)	Retrospective cancer registry review	7 breast cancer centers, Malaysia	ANN achieved 84.5% survival prediction accuracy	Supported personalized treatment stratification	Predicted 5-year survival probability	High accuracy prediction; aided decision-making	Missing data; socioeconomic factors not included; model comparison complexity

Moksem & Newaz (2024)	Prospective (Survey)	Malaysia (UNIRAZAK and UMMC)	AI automated tumor contouring, reducing manual errors	AI improved dosimetry and beam optimization	Dynamic plan adjustments during treatment	30-60% improved efficiency; high clinical acceptability	Survey bias; unexplained variance; technical limits
Zhang et al. (2024)	Retrospective (Secondary big data analysis :TCGA, SEER)	Global (with Malaysia collaboration)	99% diagnosis accuracy (multi-omics ML models)	AI stratified tumor types for therapy	Survival and recurrence predictions from multi-omics data	Highly accurate predictions; efficient data processing	Data privacy, standardization, generalizability issues
Abdul Rasool Hassan et al. (2025)	Neither retrospective not prospective	Global (Monash University Malaysia included)	AI imaging improved early cancer detection (87% sensitivity)	ML predicted chemotherapy responses	ML predicted toxicity and progression risks	20–30% improved treatment outcomes; workflow automation	Bias, clinical validation needed, ethical concerns
Aggarwal et al. (2023)	Prospective observational study	Malaysia, India, Jordan, South Africa	AI automated contouring (target/OARs)	Reduced RT planning time from weeks to minutes	Quality assurance for plan compliance	95% plan acceptability; significant cost savings	No patient outcome tracking; infrastructure variation
Lim et al. (2021)	Retrospective(Digital mammogram	UTHM	Hybrid SVM-ANN achieved 100% benign vs.	-	-	99.4% overall diagnostic accuracy; reduced	Small , dataset only mammography used

	analysis :Mini MIAS dataset)		malignant classification			radiologist workload	
Ling et al. (2022)	Retrospective (Secondary data + Literature review)	Malaysia (UTAR, Xiamen Univ. Malaysia)	AI + miRNA biomarker profiling >90% accuracy	AI predicted personalized therapy responses	Predicted recurrence, monitored progression	Less invasive; early detection; reduced cost	miRNA variability; clinical integration challenges
Mahdi et al. (2024)	Retrospective (PET/CT imaging dataset: HECKTOR Challenge)	Multinational (including Malaysia PET/CT centers)	Improved tumor segmentation using Weighted Fusion Transformer	Accurate RT planning from better tumor maps	Tumor volumetry tracked treatment response	Faster, more accurate segmentation; reduced radiologist burden	PET/CT fusion computational complexity; lymph node segmentation
Mirza et al. (2023)	Retrospective (GC image datasets :Kvasir)	Global (incl. Malaysia, USM)	Deep learning + Hybrid optimization for GC improved classification	-	Timely relapse detection of GC lesion	Higher early GC detection rates; reduced endoscopy need	Focused only on GC cancers; the model needs further testing

*Note. The table is summary of data extraction of the selected studies. Symbols like “+” indicate and, “-” indicate no result found, “>” indicate is more than, “→” indicate to and “%“is percentages. Abbreviation is also used and explained in the list of abbreviations.

ORIGINAL ARTICLE

Development and Validation of a Questionnaire on Knowledge and Attitude Toward Far-Infrared-Emitting Garments, Their Medical Applications and Health Benefits

[1]Haizatul Umiyah Nor Azmi[1], [1]Ibrahim Adham Taib[1], Nor Saadah Md Azahri[1,*]

[1]Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
Far-infrared-emitting garments; Questionnaire; Knowledge; Attitude; Validity; Reliability

ABSTRACT

Introduction: Infrared radiation is a form of energy within the electromagnetic spectrum, categorized by its frequency and wavelength. Far-infrared (FIR)-emitting garments such as blankets, belts, clothing, and socks infused with bio-ceramic powder are commercially available and have been associated with various health benefits in scientific literature. However, no studies have yet explored the public's knowledge, attitude, and practice (KAP) regarding the use of FIR-emitting garments. This highlights the need for a valid and reliable instrument to assess these components. **Objective:** The primary objective of this study was to develop and validate a questionnaire to assess knowledge and attitudes toward FIR-emitting garments, including their medical applications and health benefits. **Methodology:** The development process included both validity and reliability assessments. Five experts conducted face and content validity evaluations based on item relevance, clarity, simplicity, and ambiguity. The Content Validity Index (CVI) was calculated using both the Item-level (I-CVI) and Scale-level (S-CVI) indices. A pilot test involving 51 respondents was carried out to determine internal consistency using Cronbach's Alpha. **Results:** The CVI scores for knowledge items averaged 0.88, with with I-CVI and S-CVI values averaging 0.88 across all evaluated criteria. For attitude items, the I-CVI and S-CVI values averaged 0.84 for relevance and 0.87 for clarity, simplicity, and ambiguity. The reliability analysis yielded Cronbach's Alpha values of 0.807 for knowledge items and 0.745 for attitude items. **Conclusion:** The developed questionnaire demonstrated strong content validity (CVI > 0.83) and acceptable internal consistency (Cronbach's Alpha > 0.73). Therefore, the instrument is considered valid and reliable for assessing knowledge and attitudes related to FIR-emitting garments and their potential health benefits. It is recommended for use in future population-based studies.

INTRODUCTION

Infrared (IR) radiation, also known as thermal radiation, is part of the electromagnetic spectrum and is classified according to frequency and wavelength (Bontemps et al., 2021). The International Commission on Illumination (CIE) classifies infrared radiation into three sub-bands based on wavelength: near-infrared (0.7–1.4 µm), mid-infrared (1.4–3.0 µm), and far-infrared (3.0–100 µm) (Vatansever & Hamblin, 2012). Far-infrared (FIR) radiation, while invisible to the naked eye, is perceived by human thermoreceptors in the skin as radiant heat. It can penetrate the body to a depth of approximately 2 to 3 centimetres into subcutaneous tissue without causing excessive surface heating (Hartel et al., 2006, as cited in Shui et al., 2015).

FIR-emitting garments typically contain bioceramic materials either embedded as fibres or powders within the fabric or affixed as disks or patches. These materials, composed of metallic oxides such as magnesium, silica, aluminium, and tourmaline, are capable of emitting infrared and far-infrared radiation when stimulated by body heat (Bontemps et al., 2021). Once heated, these garments absorb body heat, retain it, and re-emit it as FIR radiation back to the body. This interaction may result in physiological responses such as elevated skin temperature, vasodilation, and enhanced microcirculation in targeted areas (Mantegazza et al., 2018).

A range of FIR-emitting products is available commercially, including blankets, belts, clothing, and socks that incorporate bioceramic powders (Dyer, 2011, as cited in Bontemps et al., 2021). Research suggests these garments may offer multiple health benefits. Reported effects include reduction of skin temperature during exercise by promoting peripheral blood flow (Kenny & McGinn, 2017

* *Corresponding author.*
E-mail address: nsaadah@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

as cited in Bontemps et al., 2021), enhanced oxygen utilization, delayed onset of anaerobic metabolism, and prolonged physical endurance (Mantegazza et al., 2018). Additional findings indicate improvements in sleep quality (McCall et al., 2018) and reductions in body measurements such as weight (Conrado & Munin, 2011).

MATERIALS AND METHODS

Ethical Approval

This study received ethical clearance from the Kulliyyah Postgraduate and Research Committee (KPGRC), Kulliyyah of Allied Health Sciences, and the IIUM Research Ethics Committee (IREC). All data collected from participants were treated with strict confidentiality and used solely for research purposes, including analysis and reporting.

Questionnaire Development

As there was no existing validated instrument assessing knowledge and attitudes regarding far-infrared (FIR)-emitting garments, a new questionnaire was developed based on a comprehensive review of scientific literature published between 1999 and 2021. Earlier studies, particularly those published before 2010, were included because a significant number of foundational experiments on the health effects of FIR-emitting garments involved animal models conducted during that period.

The questionnaire was divided into three sections:

- Section A: Sociodemographic information
- Section B: Knowledge-related items (12 items)
- Section C: Attitude-related items (11 items)

All items were close ended. To encourage thoughtful responses, a mixture of positively and negatively worded statements was used. Respondents were instructed to select only one answer per item.

The questionnaire was developed in English, and the estimated completion time was less than 15 minutes.

- In Section B (Knowledge), response options included *"True," "False,"* and *"Not Sure"*.
- In Section C (Attitude), a 5-point Likert scale was used, ranging from *"Strongly Agree" (1)* to *"Strongly Disagree" (5)*.

Scoring was conducted as follows: each correct answer in the knowledge section received one point, while incorrect or "Not Sure" responses received zero points. For attitude items, response distributions were reported in percentages for each Likert scale level.

Questionnaire Validation

Face Validity

Face validity was evaluated to ensure grammatical accuracy, logical item flow, and content appropriateness. This process followed guidance from DeVon et al. (2007), who suggested involving domain experts or laypeople to assess the instrument's clarity, structure, and relevance. A total of five experts participated in the face and content validation process. The panel consisted of academic staff from the Kulliyyah of Allied Health Sciences, specializing in:

- Anatomy and Physiology (1 expert)
- Physiology (1 expert)
- Epidemiology and Research Methodology (2 experts)
- Medical Imaging (1 expert)

A comment section was included in the validation form for each item to allow reviewers to provide feedback and suggestions for improvement.

Content Validity

Content validity refers to the extent to which the questionnaire items adequately reflect the theoretical constructs under investigation (Schultz & Whitney, 2005, as cited in Tsang et al., 2017). Experts rated each item on four criteria: relevance, clarity, simplicity, and ambiguity using a 4-point Likert scale. The Content Validity Index (CVI) was calculated, comprising:

- I-CVI (Item-Level Content Validity Index)
- S-CVI (Scale-Level Content Validity Index)

These indices quantified the proportion of agreement among the expert reviewers for each item and the overall scale.

The other calculation to determine the value of alpha (α) was measured by using:

$$\alpha = \frac{\kappa}{\kappa - 1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_x^2}\right)$$

x = a questionnaire
k = number of items
σ_i^2 = is the variance of item i
σ_x^2 = total variance of the questionnaire

Questionnaire Reliability

Reliability pertains to the consistency and repeatability of the instrument under similar conditions (Razavipour & Raji, 2022; Ofem & Mchi, 2023). Internal consistency was evaluated using Cronbach's Alpha (α), which measures the degree to which items within a scale are interrelated (Tsang et al., 2017). The reliability analysis was conducted using IBM SPSS Statistics version 26.0**.**

Before computing Cronbach's Alpha, all negatively worded items were reverse coded to ensure appropriate scoring:

- Knowledge items were scored as 1 for correct and 0 for incorrect/'Not Sure'. *Negatively phrased items were reverse-coded accordingly.*
- For attitude items, the reverse coding followed this logic:
    - *Strongly Agree (1)* → recoded as *Strongly Disagree (5)*
    - *Agree (2)* → recoded as *Disagree (4)*
    - *Neutral (3)* remained unchanged
    - *Disagree (4)* → recoded as *Agree (2)*
    - *Strongly Disagree (5)* → recoded as *Strongly Agree (1)*

According to Streiner (2003), reverse-coding is essential to prevent artificially low Cronbach's Alpha values resulting from negative item correlations. A pilot test involving 51 participants was conducted to calculate the reliability coefficients.

Interpretation of Correlation Coefficients

The interpretation of Cronbach's Alpha was guided by Mukaka (2012), as shown in **Table 1.**

Table 1: Interpretation of Correlation Coefficient Magnitudes (Mukaka, 2012)

Size of Correlation	Interpretation
0.90 to 1.00 (±)	Very high positive (negative)
0.70 to 0.90 (±)	High positive (negative)
0.50 to 0.70 (±)	Moderate positive (negative)
0.30 to 0.50 (±)	Low positive (negative)
0.00 to 0.30 (±)	Negligible correlation

Study Setting

Face and content validity evaluations were conducted remotely via email communication. Experts received a digital validation form containing clear instructions and the draft questionnaire for assessment.

The pilot test was administered to undergraduate students from the Kulliyyah of Allied Health Sciences (KAHS). The

questionnaire link was distributed via WhatsApp, both through personal messages and student group chats. The primary aim of the pilot was to determine the internal consistency and reliability of the developed instrument measuring knowledge and attitudes toward FIR-emitting garments and their health-related applications.

RESULTS

Face Validity

Knowledge Domain

Expert evaluation was conducted to assess the face validity of the initial 15 items in the knowledge domain. Feedback highlighted issues related to wording, clarity, and redundancy. Following the review, three items were removed, and several items were revised to improve clarity and comprehension. This process resulted in a reduction to 13 finalized items. For instance, technical terms were clarified with lay explanations, and grammatical structures were standardized to enhance readability. A summary of the changes, including original and revised items as well as the rationale for modification, is presented in Table 2.

Attitude Domain

Similarly, 15 initial items assessing attitude were reviewed. Expert feedback identified issues of redundancy, irrelevance, and inappropriate phrasing. As a result, four items were removed, and the remaining were revised for content clarity and relevance, leading to a final total of 11 attitude items. Notable adjustments included rephrasing items for consistency in tone (e.g., positively or negatively worded statements), removal of knowledge-based content, and correction of grammatical inconsistencies. The revised attitude items and corresponding changes are summarized in Table 3.

Content Validity Index (CVI)

Table 4 shows the relevance of items based on knowledge domain. Content validity was assessed quantitatively using the Item-Level Content Validity Index (I-CVI), Scale-Level Content Validity Index based on average agreement (S-CVI/Ave), and Universal Agreement (S-CVI/UA), as described by Roy & Sahu (2024). Five experts independently rated each item on a 4-point scale in terms of relevance, clarity, simplicity, and ambiguity.

Ratings of 3 or 4 were recoded as "1" (valid), while ratings of 1 or 2 were recoded as "0" (not valid). The I-CVI for each item was computed by dividing the number of experts rating an item as 3 or 4 by the total number of experts. For example, if all five experts rated an item as relevant (score 3 or 4), the I-CVI was 1.00. S-CVI/Ave was derived by averaging the I-CVI scores across all items, while S-CVI/UA reflected the proportion of items that achieved full agreement across all experts.

For knowledge items, the S-CVI/Ave was 0.88, indicating strong overall content validity, while the S-CVI/UA was 0.67. A summary of CVI values for relevance, clarity, simplicity, and ambiguity for knowledge items is presented in Table 5.0.

For attitude items, the S-CVI/Ave values were 0.84 for relevance and 0.87 for clarity, simplicity, and ambiguity. The S-CVI/UA for attitude items was 0.67 (see **Table 6.0**). Two attitude items were excluded due to conceptual overlap with knowledge items or low relevance scores.

Tables summarizing the CVI calculations for relevance (Table 4.0), and comprehensive breakdowns for each item's CVI dimensions (Tables 5.0 and 6.0), are provided.

Table 2: The List of Updated Items for Development of Questionnaire on Knowledge on Far-Infrared-Emitting Garments, Their Medical Applications and Health Benefits

Draft of Items	Updated Items	Action for Improvement Based on Feedbacks
(Q1) Infrared radiation is one of the region in electromagnetic spectrum	Thermal radiation (infrared) is one of the regions of the electromagnetic spectrum (range of all types of electromagnetic radiation)	Add some definition and use layman term (put in bracket after the original term) to make sure the respondents understood the question
(Q2) Infrared radiation can be divided into 3 groups which are near infrared, mid infrared and far infrared	Infrared radiation can be divided into near-infrared, mid-infrared and far-infrared	Change the structure of item
(Q3) Far infrared radiation is visible #	Far-infrared radiation is visible #	-
(Q4) The wavelength of far infrared radiation is 3.0 µm to 100 µm	The wavelength of far-infrared radiation is 3.0 µm to 100 µm	Correct the grammatical error
(Q5) Far infrared cannot penetrate the skin #	Far-infrared cannot penetrate the skin #	-
(Q6) Far infrared radiation can cause dilation of capillary	Far-infrared radiation can promote dilation of blood capillaries	Change the structure of item
(Q7) Far infrared emitting garments are type of items made of bio ceramic fibres	Garments that emit far-infrared are made up of bioceramic fibres	Correct the grammatical errors
(Q8) The materials of bio ceramic fibres can radiate far infrared when heated to body temperature	When heated to body temperature, bioceramic fibres can emit far-infrared	Change the structure of the item
(Q9) Far infrared emitting garments cannot reduce skin temperature during exercise #	Far-infrared-emitting garments can reduce skin temperature during exercise	Putting as positively worded item
(Q10) Far infrared emitting socks act as cooling agent during long distance running exercise	-	Delete the item because contradict with item number 9
(Q11) Far infrared emitting garments enhancing oxygen production	Far-infrared-emitting garments can enhance oxygen production	Correct the grammatical errors
(Q12) Far infrared emitting garments can delay anaerobic metabolism during exercise	Far-infrared-emitting garments can delay the onset of fatigue during exercise	Change the structure of the item
(Q13) Far infrared bedsheet can promote sleep induction	Far-infrared-emitting bedsheets can reduce sleep quality #	Change to negative worded item and change the structure of the item

Note. Negative items are identified with (#)

Table 3: The List of Updated Items for Development of Questionnaire on Attitude on Far-Infrared-Emitting Garments, Their Medical Applications and Health Benefits

Draft of items	Updated items	Action for improvement based on feedbacks
(Q1) I believe far infrared emitting garments can provide many health benefits	I believe that far-infrared-emitting garments can provide many health benefits	Correct the grammatical error
(Q2) I think far infrared emitting garments can cause burn injury to skin **#**	I fear far-infrared-emitting garments can burn the skin #	Correct the grammatical error
(Q3) I feel far infrared radiation cannot penetrate the skin due to its wavelength #	-	Delete the item because repetition in Q5 in knowledge domain
(Q4) I am aware with the manufactured of far infrared emitting garments can be found in clothes	I am aware of the manufacturing of far-infrared-emitting clothes	Correct the grammatical error and the structure of the item
(Q5) I think far infrared emitting garments cannot reduce skin temperature during exercise #	-	Delete the item because repetition in Q9 in knowledge domain
(Q6) I believed far infrared emitting garments can longer the duration of exercise	I do not believe far-infrared-emitting garments can reduce fatigue during exercise #	Change the structure of item and putting as negative worded item
(Q7) I am aware with far infrared emitting socks can reduce unpleasant smells during exercise	I believe that far-infrared-emitting socks can reduce unpleasant smells during exercise	Correct the grammatical error
(Q8) I feel the usage of far infrared emitting garments can reduce comfort during exercise #	I feel wearing far-infrared-emitting garments can be uncomfortable #	Change the structure of item
(Q9) I am aware the manufacturer of far infrared emitting garments can be found in belts	-	Delete the item because sound like question on knowledge
(Q10) I believe far infrared belts can be used for weight reduction	I believe that far-infrared-emitting belts can be used for weight reduction	Correct the grammatical error
(Q11) I am aware far infrared emitting garments can reduce menstrual pain	I believe that far-infrared-emitting garments can help reduce menstrual pain	Correct the grammatical error
(Q12) I think far infrared bed sheet can improve sleep quality	I think far-infrared-emitting bedsheets can be used to treat insomnia	Change the structure of item
(Q13) I believe the manufacturer of far infrared in sportswear can improve performance in athletes	I believe far-infrared-emitting sportswear can improve performance in athletes	Change the structure of item
(Q14) I believe far infrared emitting garments can prevent hyperthermia during hot temperature	I believe far-infrared-emitting garments can cause heat stroke during hot weather #	Change the structure of item and putting as negative worded item
(Q15) I think far infrared emitting garments can optimise the recovery process in athletes	-	Delete the item because the question sound like knowledge

Note. Negative items are identified with (#)

Internal Consistency Reliability

Knowledge Domain

Internal consistency reliability for the knowledge domain was assessed using Cronbach's Alpha. Table 7 shows the mean score of each item on knowledge domain. Total of 12 items were initially analysed using data from 51 respondents. The inter-item correlation matrix revealed weak correlations ($r < 0.3$) for several item pairs (Table 8), and item-total statistics (Table 9) showed that the removal of item K1 improved the internal consistency. After removing this item, the Cronbach's Alpha increased from 0.795 to 0.807 (Table 10), indicating good internal consistency. The final knowledge domain contained 11 items.

Attitude Domain

Table 11 shows mean score of each item for attitude domain. There was no item with mean higher than range from 2.61 to 3.29. Table 12 shows the value of inter-item correlation matrix and Table 13 shows the value of corrected items and value of Cronbach's Alpha if item deleted.

The internal consistency of the attitude domain was similarly assessed using Cronbach's Alpha. Initial analysis of 11 items revealed two problematic items: A3 and A11. These items demonstrated low item-total correlations and negatively affected the overall reliability. After their removal, the Cronbach's Alpha increased to 0.729, confirming acceptable internal consistency (Table 14). The final attitude scale consisted of 9 items.

Summary of Modifications and Psychometric Properties

- Total items in the knowledge domain were reduced from 15 to 11 after revisions based on expert feedback and reliability analysis.
- Total items in the attitude domain were reduced from 15 to 9.
- All retained items demonstrated acceptable to excellent CVI values.
- Cronbach's Alpha values indicated satisfactory internal consistency for both domains (knowledge = 0.807; attitude = 0.729).

Table 4: The Calculation of Content Validity Index for Relevance on Knowledge Items

Domain : Knowledge	Relevance							
Items	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert in agreement	I-CVI	UA
1	1	1	1	1	1	5	1	1
2	1	1	1	1	1	5	1	1
3	1	1	1	1	1	5	1	1
4	1	1	1	1	1	5	1	1
5	1	1	1	1	1	5	1	1
6	1	1	1	1	1	5	1	1
7	1	1	1	1	1	5	1	1
8	1	1	1	1	1	5	1	1
9	1	0	1	1	1	4	0.8	0
10	1	0	1	1	1	4	0.8	0
11	1	1	1	1	1	5	1	1
12	0	1	1	1	1	4	0.8	0
13	1	1	1	1	1	5	1	1
14	0	0	1	0	0	1	0.2	0
15	0	0	1	1	1	3	0.6	0
Proportion relevance =	0.8	0.73	1	0.93	0.93	S-CVI Ave =	0.88	
						S-CVI/UA =		0.67
Average proportion of item judged as relevance across the 5 experts =				0.88				

Table 5: The Summary of Content Validity Index (CVI) based on Knowledge Items for Relevance, Clarity, Simplicity and Ambiguity

Items	Relevance	Clarity	Simplicity	Ambiguity
(Q1) Infrared radiation is one of the region in electromagnetic spectrum	1.0	0.8	0.8	0.8
(Q2) Infrared radiation can be divided into 3 groups which are near infrared, mid infrared and far infrared	1.0	1.0	0.8	1.0
(Q3) Far infrared radiation is visible	1.0	1.0	1.0	1.0
(Q4) The wavelength of far infrared radiation is 3.0 µm to 100 µm	1.0	1.0	1.0	1.0
(Q5) Far infrared cannot penetrate the skin	1.0	1.0	1.0	1.0
(Q6) Far infrared radiation can cause dilation of capillary	1.0	1.0	1.0	1.0
(Q7) Far infrared emitting garments are type of items made of bio ceramic fibres	1.0	0.8	1.0	0.8
(Q8) The materials of bio ceramic fibres can radiate far infrared when heated to body temperature	1.0	1.0	1.0	1.0

(Q9) Far infrared emitting garments cannot reduce skin temperature during exercise	0.8	1.0	1.0	1.0
(Q10) Far infrared emitting socks act as cooling agent during long distance running exercise	0.8	1.0	1.0	1.0
(Q11) Far infrared emitting garments enhancing oxygen production	1.0	0.6	0.8	0.6
(Q12) Far infrared emitting garments can delay anaerobic metabolism during exercise	0.8	0.8	0.8	0.8
(Q13) Far infrared bedsheet can promote sleep induction	1.0	1.0	0.8	1.0
(Q14) The factors of sleepiness are depending on the accumulation of extracellular adenosine	0.2	0.4	0.4	0.4
(Q15) Far infrared bedsheet cannot decline the accumulation of extracellular adenosine level	0.6	0.8	0.8	0.8
(S-CVI/Ave) based on (I-CVI)	0.88	0.88	0.88	0.88
(S-CVI/UA)	0.67	0.6	0.53	0.6
(S-CVI/Ave) based on proportion	0.88	0.88	0.88	0.88

Table 6: The Summary of Content Validity Index (CVI) based on Attitude Items for Relevance, Clarity, Simplicity and Ambiguity

Items	**Relevance**	**Clarity**	**Simplicity**	**Ambiguity**
(Q1) I believe far infrared emitting garments can provide many health benefits	1.0	1.0	1.0	1.0
(Q2) I think far infrared emitting garments can cause burn injury to skin	0.8	1.0	1.0	1.0
(Q3) I feel far infrared radiation cannot penetrate the skin due to its wavelength	0.4	0.8	0.8	0.6
(Q4) I am aware with the manufactured of far infrared emitting garments can be found in clothes	0.8	0.4	0.4	0.4
(Q5) I think far infrared emitting garments cannot reduce skin temperature during exercise	0.4	1.0	1.0	1.0
(Q6) I believed far infrared emitting garments can longer the duration of exercise	1.0	1.0	1.0	1.0
(Q7) I am aware with far infrared emitting socks can reduce unpleasant smells during exercise	1.0	1.0	1.0	1.0
(Q8) I feel the usage of far infrared emitting garments can reduce comfort during exercise	0.8	0.8	0.8	0.8
(Q9) I am aware the manufacturer of far infrared emitting garments can be found in belts	0.4	0.2	0.2	0.4
(Q10) I believe far infrared belts can be used for weight reduction	1.0	1.0	1.0	1.0
(Q11) I am aware far infrared emitting garments can reduce menstrual pain	1.0	1.0	1.0	1.0
(Q12) I think far infrared bed sheet can improve sleep quality	1.0	1.0	1.0	1.0
(Q13) I believe the manufacturer of far infrared in sportswear can improve performance in athletes	1.0	0.8	0.8	0.8
(Q14) I believe far infrared emitting garments can prevent hyperthermia during hot temperature	1.0	1.0	1.0	1.0
(Q15) I think far infrared emitting garments can optimise the recovery process in athletes	1.0	1.0	1.0	1.0
(S-CVI/Ave) based on (I-CVI)	0.84	0.87	0.87	0.87
(S-CVI/UA)	0.67	0.67	0.67	0.67
(S-CVI/Ave) based on proportion	0.84	0.87	0.87	0.87

Table 7: The Value of Means of Each Items on Knowledge Items.

Item Statistics			
	Mean	Std. Deviation	N
K1	.86	.348	51
K2	.63	.488	51
K3#	.24	.428	51
K4	.16	.367	51
K5#	.27	.451	51
K6	.37	.488	51
K7	.27	.451	51
K8	.25	.440	51
K9	.20	.401	51
K10	.20	.401	51
K11	.31	.469	51
K12#	.12	.325	51

Table 8: The Value of Inter-Item Correlation Matrix on Knowledge Items

Inter-Item Correlation Matrix												
	K1	K2	K3#	K4	K5#	K6	K7	K8	K9	K10	K11	K12#
K1	1.000	.164	.087	.172	-.010	.072	-.010	-.028	-.090	-.090	.024	.146
K2		1.000	.236	.109	.110	.174	.201	.172	.176	.074	.259	.030
K3#			1.000	-.112	.073	.242	.177	.100	.192	.075	.123	.228
K4				1.000	.339	.225	.218	.243	.194	.330	.289	-.157
K5#					1.000	.344	.212	.245	.028	.028	.152	.048
K6						1.000	.526	.573	.539	.641	.440	.096
K7							1.000	.648	.471	.582	.626	.321
K8								1.000	.391	.504	.671	.205
K9									1.000	.751	.518	-.027
K10										1.000	.518	-.027
K11											1.000	.409
K12#												1.000

Table 9: The Value of Corrected Item and Cronbach's Alpha if Item Deleted on Knowledge Items

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
K1	3.02	7.660	.065	.172	.807
K2	3.25	6.954	.277	.180	.796
K3#	3.65	7.193	.230	.196	.797
K4	3.73	7.163	.309	.382	.789
K5#	3.61	7.083	.258	.341	.796
K6	3.51	6.055	.671	.619	.750
K7	3.61	6.163	.688	.588	.750
K8	3.63	6.278	.650	.610	.755
K9	3.69	6.620	.545	.640	.767
K10	3.69	6.540	.588	.746	.763
K11	3.57	6.090	.690	.679	.749
K12#	3.76	7.424	.212	.429	.795

Table 10: The Value of Cronbach's Alpha after Item Number 1 Deleted on Knowledge Items

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.807	.801	11

Table 11: The Value of Means of each Items on Attitude Items.

Item Statistics

	Mean	Std. Deviation	N
A1	2.61	.918	51
A2#	3.08	.913	51
A3	3.29	1.137	51
A4#	2.78	.673	51
A5	2.75	.891	51
A6#	3.06	.835	51
A7	3.08	.771	51
A8	3.00	.800	51

A9	2.90	.855	51
A10	2.90	.831	51
A11#	3.08	.845	51

Table 12: The Value of Inter-Item Correlation Matrix on Attitude Items

Inter-Item Correlation Matrix

	A1	A2#	A3	A4#	A5	A6#	A7	A8	A9	A10	A11#
A1	1.000	-.153	.055	.216	.413	-.021	.384	.572	.485	.552	-.243
A2#		1.000	.074	-.037	-.098	.073	.020	.137	-.041	.221	.070
A3			1.000	.215	.174	-.103	.064	.286	-.031	-.011	-.025
A4#				1.000	.374	.023	.342	.297	.206	.391	-.005
A5					1.000	-.168	.292	.196	.334	.344	-.159
A6#						1.000	-.007	-.030	-.048	-.107	.334
A7							1.000	.616	.498	.606	-.225
A8								1.000	.497	.602	-.118
A9									1.000	.578	-.543
A10										1.000	-.359
A11#											1.000

Table 13: The Value of Corrected Item and The Value of Cronbach's Alpha If Item Deleted on Attitude Items

Item-Total Statistic

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
A1	29.92	15.074	.474	.558	.563
A2#	29.45	18.053	.053	.269	.653
A3	29.24	16.744	.135	.282	.647
A4#	29.75	16.474	.434	.316	.583
A5	29.78	15.973	.355	.373	.590
A6#	29.47	18.734	-.019	.208	.661
A7	29.45	15.373	.551	.523	.556
A8	29.53	14.454	.690	.680	.524

A9	29.63	15.878	.394	.591	.583
A10	29.63	14.758	.602	.676	.540
A11#	29.45	20.413	-.240	.477	.699

Table 14: The Value of Cronbach's Alpha after Item Number 11 Deleted on Attitude Items

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.729	.737	9

DISCUSSION

Analysis of Face Validity and Content Validity Index (CVI) for Relevance, Clarity, Simplicity, and Ambiguity on Knowledge Items

Table 5 summarizes CVI values for relevance, clarity, simplicity, and ambiguity of items within the knowledge domain. All items scored above 0.8 for item-level CVI (I-CVI), except for item 11 (clarity = 0.6, ambiguity = 0.6), item 14 (relevance = 0.2, clarity = 0.4, simplicity = 0.4, ambiguity = 0.4), and item 15 (relevance = 0.6). According to Polit & Beck (2006), items evaluated by five or fewer experts must achieve a CVI of at least 0.83 to be considered valid. Thus, items scoring below 0.83 did not meet the threshold for content validity.

Based on expert feedback from face validity assessments, item 14 ("The factors of sleepiness are depending on the accumulation of extracellular adenosine") was removed as it did not reflect the knowledge of far-infrared-emitting garments, their medical applications, or health benefits. Item 15 ("Far-infrared bedsheets cannot decline the accumulation of extracellular adenosine level") was also deleted due to redundancy and contradiction with item 14.

Item 11 was retained but revised for grammatical clarity. The phrase "Far infrared emitting garments" was corrected to "Far-infrared-emitting garments," and "enhancing" was revised to "can enhance" oxygen production. As stated by Polit & Beck (2006), items should be revised or eliminated if they do not meet the minimum agreement among experts.

The average S-CVI/Ave (scale-level CVI based on averaging I-CVI scores across items) for relevance, clarity, simplicity, and ambiguity was 0.88, exceeding the 0.83 threshold required for acceptable content validity (Lynn, 1986). However, the S-CVI/UA (scale-level CVI based on universal agreement) values were lower: relevance (0.67), clarity (0.6), simplicity (0.53), and ambiguity (0.6). These scores did not meet the 100% agreement criterion for universal agreement. Polit and Beck (2006) noted that as the number of experts increases, the likelihood of total agreement tends to decrease.

Analysis of Face Validity and Content Validity Index (CVI) for Relevance, Clarity, Simplicity, and Ambiguity on Attitude Items

Table 6 displays CVI values for the attitude items. Items with I-CVI values below 0.8 included item 3 (relevance = 0.4, ambiguity = 0.6), item 4 (clarity = 0.4, simplicity = 0.4, ambiguity = 0.4), item 5 (relevance = 0.4), and item 9 (relevance = 0.4, clarity = 0.2, simplicity = 0.2, ambiguity = 0.4). These values did not meet the minimum required I-CVI of 0.83 (Shen et al. 2024).

Item 3 ("I feel far infrared radiation cannot penetrate the skin due to its wavelength") was deleted due to redundancy with item 5. Likewise, item 5 was deleted due to overlap with item 9. Item 9 ("I am aware the manufacturer of far infrared emitting garments can be found in belts") was also removed as it resembled a knowledge-based question. Item 4 was revised for grammar and clarity, changing from "I am aware with the manufactured of far infrared emitting garments can be found in clothes" to "I am aware of the manufacturing of far-infrared-emitting clothes."

The S-CVI/Ave across all attitude items was 0.87, except for relevance which scored 0.84—both above the minimum threshold of 0.83 for five or fewer experts (Ma et al. 2022). The S-CVI/UA values were 0.67 for clarity, simplicity, and ambiguity, and 0.6 for relevance, all below the required 1.0 (100%) agreement. This reflects the challenge of achieving

full agreement among all experts due to different background of experts.

Analysis of Questionnaire Reliability Based on Cronbach's Alpha for Knowledge Items

Table 7 shows the mean scores for each knowledge item, which ranged from 0.20 to 0.37, indicating consistency among items. However, items 1 (mean = 0.86) and 2 (mean = 0.63) were outliers.

Table 8 indicates that item 1 had negligible correlations (below 0.3) with other items, along with five negative inter-item correlations. According to Mukaka (2012), such low correlation suggests the item does not measure the same construct and should be considered for removal. Item 1 was therefore removed.

Table 9 shows that item 1 had the lowest corrected item-total correlation (0.065) and the highest increase in Cronbach's Alpha when deleted (0.807). Experts also suggested improving clarity by using layman's terms. The item was revised from "Infrared radiation is one of the regions in the electromagnetic spectrum" to "Thermal radiation (infrared) is one of the regions of the electromagnetic spectrum (range of all types of electromagnetic radiation)."

After item 1 was deleted, the Cronbach's Alpha increased to 0.807 (Table 10.0), which is within the acceptable range of 0.70–0.90 for high reliability (Mukaka, 2012). The number of knowledge items was reduced from 12 to 11.

Analysis of Questionnaire Reliability Based on Cronbach's Alpha for Attitude Items

Table 11 indicates that the mean scores of attitude items ranged from 2.61 to 3.29, with no extreme outliers. Table 12.0 reveals that items 6 and 11 had the most negative inter-item correlations, likely due to being reverse-coded. Their mixed correlations suggested these items were problematic.

Table 13 shows that deleting item 11 would increase Cronbach's Alpha to 0.699, followed by items 6 (0.661), 2 (0.653), and 3 (0.647). Although Taber (2018) recommends a minimum alpha of 0.70, item 3 was deleted based on both expert feedback and low CVI values (relevance = 0.8, clarity = 0.4, simplicity = 0.4, ambiguity = 0.4).

After deleting item 3, Cronbach's Alpha was 0.647, so item 11 was also removed to achieve an acceptable reliability level. The final Cronbach's Alpha (Table 14) met the threshold of ≥0.70, and the number of attitude items was reduced to 9. Item 11 was revised from "I believe far infrared emitting garments can prevent hyperthermia during hot temperature" to "I believe far-infrared-emitting garments can cause heat stroke during hot weather," as recommended by experts for improved clarity.

CONCLUSION

This study demonstrated that the developed questionnaire is a valid and reliable tool for assessing knowledge and attitudes towards far-infrared-emitting garments, their medical applications, and health benefits. Both the CVI and Cronbach's Alpha values met the required thresholds for five or fewer experts and internal consistency, respectively.

However, some limitations were identified. The presence of negatively worded or reverse-coded items in the knowledge and attitude sections may have reduced reliability. This contradicts item development guidelines, which recommend avoiding negative formulations (Haladyna, Downing & Rodríguez, 2002; Haladyna & Rodríguez, 2013; Lane, Raymond & Haladyna, 2016; Moreno et al., 2004, 2006, 2015).

Additionally, the use of a 5-point Likert scale may have affected Cronbach's Alpha. Polit and Beck (2006) suggest using a 4-point scale to eliminate neutral or ambivalent midpoints. Future studies should consider developing a revised questionnaire using only positively worded items and a 4-point scale to further improve reliability and validity.

ACKNOWLEDGEMENT

This research was not funded by any grant. The author would like to express sincere gratitude to the lecturers of the Kulliyyah of Allied Health Sciences for their support and contributions of expertise throughout this study. Special thanks also go to the respondents for their time and willingness to participate in the survey.

REFERENCES

Bontemps, B., Gruet, M., Vercruyssen, F., & Louis, J. (2021). Utilisation of far infrared-emitting garments for optimising performance and recovery in sport: Real potential or new fad? PLOS ONE, 16(5), e0251282. https://doi.org/10.1371/journal.pone.0251282

Conrado, L. A., & Munin, E. (2011). Reduction in body measurements after use of a garment made with synthetic fibers embedded with ceramic nanoparticles.

Journal of Cosmetic Dermatology, 10(1), 30–35. https://doi.org/10.1111/j.1473-2165.2010.00537.x

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S. M., & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. Journal of Nursing Scholarship, 39(2), 155–164. https://doi.org/10.1111/j.1547-5069.2007.00161.x

Dyer, J. (2011). Functional textiles for improved performance, protection and health: Infrared functional textiles. In N. Pan & G. Sun (Eds.), Functional textiles for improved performance, protection and health (pp. 184–197). Woodhead Publishing.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. Applied Measurement in Education, 15(3), 309–334. https://doi.org/10.1207/S15324818AME1503_5

Haladyna, T. M., & Rodriguez, M. C. (2013). Developing and validating test items (1st ed.). Routledge. https://doi.org/10.4324/9780203850381

Hartel, M., Hoffmann, G., Wente, M. N., Martignoni, M. E., Büchler, M. W., & Friess, H. (2006). Randomized clinical trial of the influence of local water-filtered infrared A irradiation on wound healing after abdominal surgery. The British Journal of Surgery, 93(8), 952–960. https://doi.org/10.1002/bjs.5429

Kenny, G. P., & McGinn, R. (2017). Restoration of thermoregulation after exercise. Journal of Applied Physiology, 122(4), 933–944.

Lane, S., Raymond, M. R., Haladyna, T. M., & Downing, S. M. (2016). Test development process. In S. Lane, M. R. Raymond, & T. M. Haladyna (Eds.), Handbook of test development (2nd ed., pp. 3–18). Taylor & Francis.

Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35(6), 382–385.

Ma, G., Zhong, Z., Duan, Y., Shen, Z., Qin, N., & Hu, D. (2022). *Development and Validation of a Self-Quantification Scale for Patients With Hypertension. Frontiers in Public Health, 10*, Article 849859. https://doi.org/10.3389/fpubh.2022.849859

Mantegazza, V., Contini, M., Botti, M., Ferri, A., Dotti, F., Berardi, P., & Agostoni, P. (2018). Improvement in exercise capacity and delayed anaerobic metabolism induced by far-infrared-emitting garments in active healthy subjects: A pilot study. European Journal of Preventive Cardiology, 25(16), 1744–1751. https://doi.org/10.1177/2047487318768598

McCall, W. V., Letton, A., Lundeen, J., Case, D., & Cidral-Filho, F. J. (2018). The effect of far-infrared emitting sheets on sleep. Research Journal of Textile and Apparel, 22(3), 247–259. https://doi.org/10.1108/RJTA-02-2018-0008

Moreno, R., Martínez, R., & Muñiz, J. (2004). Guidelines for the construction of multiple choice test items. Psicothema, 16(3), 490–497.

Moreno, R., Martínez, R. J., & Muñiz, J. (2006). New guidelines for developing multiple-choice items. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 2(2), 65–72. https://doi.org/10.1027/1614-2241.2.2.65

Moreno, R., Martínez, R., & Muñiz, J. (2015). Guidelines based on validity criteria for the development of multiple choice items. Psicothema.

Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Medical Journal, 24(3), 69–71.

Ofem, B. I. & Mchi, A. A. (2023). "Variable Conceptualisation and Measurement in Environmental Research". *International Journal of Methodology, 2*(1), 2–11.

Razavipour, K. & Raji, B. (2022). *Reliability of measuring constructs in applied linguistics research: a comparative study … Language Testing in Asia, 12*, Article 16.

Roy, P., & Sahu, K. K. (2024). *Building a roadmap for content validation: 6C model and validity index. International Journal of Research & Management*, 6(1), 110–114. DOI: 10.33545/26648792.2024.v6.i1b.130

Shen, Y., Zhao, M., Zhao, W., Dai, S., Wang, X., & Zhuo, Y. (2024). *Development and validation of a questionnaire to evaluate the productive engagement of Chinese older adults in the community.* BMC Public Health, 24*,* Article 3074.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique

and recommendations. Research in Nursing & Health, 29(5), 489–497. https://doi.org/10.1002/nur.20147

Shui, S., Wang, X., Chiang, J. Y., & Zheng, L. (2015). Far-infrared therapy for cardiovascular, autoimmune, and other chronic health problems: A systematic review. Experimental Biology and Medicine, 240(10), 1257–1265. https://doi.org/10.1177/1535370215573391

Shultz, K. S., & Whitney, D. J. (2005). Measurement theory in action: Case studies and exercises. SAGE Publications. https://doi.org/10.4135/9781452224749

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. Journal of Personality Assessment, 80(1), 99–103. https://doi.org/10.1207/S15327752JPA8001_18

Taber, K. S. (2018). *The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi Journal of Anaesthesia, 11(Suppl 1), S80–S89. https://doi.org/10.4103/sja.SJA_203_17

Vatansever, F., & Hamblin, M. R. (2012). Far infrared radiation (FIR): Its biological effects and medical applications. Photonics & Lasers in Medicine, 4, 255–266. https://doi.org/10.1515/plm-2012-0034

ORIGINAL ARTICLE

Traditional Malay Remedies for Women's Health: Transliteration and Scientific Insights from Malay Medical Manuscript MSS 1796

Ainul Safiyyah Zainur Rijal[1] and Mardhiah Mohammad[1,*]

[1]Department of Biomedical Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia

Keywords:
Women' health; Traditional Medicine; remedies; postpartum disorder; Malay medical manuscript

ABSTRACT

Background: The Malay medical manuscript *MSS 1796*, preserved at the Malay National Manuscript Centre, documents extensive therapeutic knowledge that reflects the intellectual heritage of the Malay world. Among its diverse medical prescription, remedies pertaining to women's health are notably prominent yet remain underexplored. **Objective:** This study examines women's health-related remedies in *MSS 1796*, focusing on three key conditions: postpartum disorders (*meroyan*), uterine prolapse (*menyacakkan peranakan*), and abdominal pain (*rengat* or *sakit ari-arinya*). **Methods:** Seventeen pages of *MSS 1796* were transliterated using standard transliteration practices by incorporating amendments to the text through corrected spellings and the addition of punctuation. Brackets "(…)" were used to indicate adjusted spellings while illustrations inserted using the phrase *"ini rajahnya […]"*, accompanied by the original images. Subsequently, the materia medica mentioned in the remedies were cross-referenced with studies sourced from PubMed and Google Scholar to assess their scientific relevance. **Results:** Fifteen remedies related to the key conditions were identified along with six other conditions. The remedies consisted of both single- and multi-ingredient, indicating the sophistication of traditional Malay pharmacology. *Oryza sativa* was frequently cited for postpartum recovery and uterine health, with scientific evidence supporting its antioxidant and anti-inflammatory properties. Other plants identified included *Curcuma longa* (turmeric) and *Trachyspermum roxburghianum* (cumin), both of which have pharmacological support for antioxidant activity and abdominal pain relief. **Conclusion:** This study highlights the significance of *MSS 1796* in preserving therapeutic knowledge related to women's health. By integrating transliteration with published pharmacological information, the study demonstrates the potential of Malay medical heritage to contribute to contemporary approaches in women's health management.

INTRODUCTION

A Malay manuscript refers to written works in the Malay language produced on readily available materials and encompassing a wide range of knowledge and wisdom (Wan Ali, 2017). They are kept by various institutions such as the National Library of Malaysia, the National Archive, the Malay National Manuscript Centre (*Pusat Kebangsaan Manuskrip Melayu*), the Malaysian Institute for Language and Literature (*Dewan Bahasa dan Pustaka*), state museums as well as abroad. Although around 22,000 Malay manuscripts are estimated to have existed, only 5,252 are known to have been collected and kept safely at the National Library of Malaysia (Hadi, 2024). In addition, a Malay medical manuscript specifically contains knowledge related to Malay medicine, traditional health practices, and therapeutic treatments practiced by Malay ancestors that are commonly known as *Kitab Tib*. (Harun Mat Piah, 2006).

The Malay Medical manuscript MSS 1796 is tagged as *Kitab Tib* by the authoritative at *Malay National Manuscript Centre* because of its medical knowledge that exceeds 80% of the content. This manuscript consists of 149 pages. According to Perpustakaan Negara Malaysia (2024), MSS 1796 is a manuscript written on European laid paper and contains formulations for treating headaches, eye diseases, and various other ailments including yaws. Moreover, the manuscript includes neither a colophon nor a title, limiting further insight into its author and its original context. Nevertheless, the nonexistence of the author's profile is common for many Malay medical manuscripts either the causes were due to missing pages or the Malay authors preferred to remain unknown. The Malay authors have been taught to abide by the ethical values of humility and sincerity while writing as the knowledge comes from the Creator and not solely from their intelligence (Mohd Affendi, 2021).

* *Corresponding author.*
E-mail address: mmoh@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

In the era of globalisation, the use of natural ingredients has grown considerably, supported by the belief that they are safer, more sustainable, and more aligned with traditional healing philosophies (Kumar et al., 2021). Across communities, there is a renewed interest in ethnomedicine, with many individuals turning to natural-based therapies despite the limited scientific validation available for certain herbal remedies. Historically, plant-based ingredients and traditional formulations have played a central role in women's health across all life stages, from adolescence and the reproductive years to menopause. Ethnomedical practices commonly utilise medicinal herbs to relieve menstrual pain (dysmenorrhea), regulate hormonal balance, support fertility, prevent miscarriage, and manage menopausal symptoms such as vasomotor instability, mood changes, and sleep disturbances (Su, 2022; Takahashi & Johnson, 2015).

Women's health is influenced by a range of conditions, including premenstrual syndrome (PMS), polycystic ovarian syndrome (PCOS), endometriosis, infertility, and postmenopausal disorders. Many of these conditions are associated with hormonal imbalances, inflammation, or oxidative stress, which can be managed or mitigated through the bioactive compounds found in medicinal plants (Su, 2022; Geller et al., 2025). For example, phytoestrogens derived from soy have been shown to mimic estrogenic activity and reduce menopausal symptoms, while herbs such as *Vitex agnus-cactus* (chasteberry) are known to regulate menstrual cycles and alleviate PMS (Höller et al., 2024; Luan et al., 2025). Additionally, *Curcuma longa* (turmeric) and *Zingiber officinale* (ginger) possess anti-inflammatory and antioxidant properties that may help reduce pelvic pain and inflammation in disorders like endometriosis (Vallée and Lecarpentier, 2020; Zhou et al., 2022).

Therefore, research on natural ingredients is crucial for identifying safe and effective treatments for women's health-related conditions. As conventional hormonal therapies and pain-management medications often carry undesirable side effects, natural alternatives may offer complementary or preventive solutions that align with holistic and personalised approaches to healthcare. This strengthens the need to re-examine Malay medical manuscripts, which preserve a rich repository of ethnomedical knowledge and diverse formulations used historically to treat women's ailments. These manuscripts provide valuable insight into traditional therapeutic strategies that may inform modern research. For example, a promising area of ongoing preclinical and clinical investigation involves plant-derived compounds and dietary interventions for the management of endometriosis, aimed at reducing inflammation, inhibiting abnormal tissue growth, and improving patients' quality of life (Meresman et al., 2021).

MATERIALS AND METHODS

Study Design

This study employed a qualitative, descriptive research design, combining library research, transliteration, and content analysis. The approach was supported by a systematic cross-referencing of traditional formulations with contemporary literature to assess their scientific relevance.

Manuscript Selection

The Malay medical manuscript *MSS 1796*, one of the *Kitab Tib* collections housed at Malay National Manuscript Centre, was selected for analysis. The manuscript is in good physical condition, with legible text and complete pagination, making it suitable for transliteration and study. Previous research had covered pages 1–40; this study extended the analysis to pages 41–57, which contained multiple formulations related to women's health.

Transliteration Process

Seventeen consecutive pages were transliterated from *Jawi* into *Rumi* script using established transliteration guidelines provided by Malaysian Institute for Language and Literature. Each *Jawi* character was systematically paired with its Roman equivalent, and adjustments were made to reflect current Malay spelling conventions. Brackets "(…)" were used to denote modernised spellings, while original spellings were retained in the appendix. Ambiguous or illegible terms were recorded as image snippets from the manuscript. References, including *A Vocabulary of Malay Medical Terms* (P.N. Gerrard), the Malay Literature Reference Centre (*Pusat Rujukan Persuratan Melayu*) database, and prior transliterations of Malay medical manuscripts, were consulted to confirm terminology.

Identification of Women's Health Formulations

Of 78 remedies documented in the transliterated text from 17 pages of the MSS 1796, 15 were identified as directly related to women's health. Emphasis was placed on remedies for postpartum disorders ('meroyan'), uterine prolapse ('menyacakkan peranakan'), postpartum abdominal pain ('rengat atau sakit ari-arinya') and any that related to women. Each formulation was catalogued using the format "1796.X.Y," where X represents the disease and Y the specific formulation number.

Content Analysis and Materia Medica Identification

The transliterated text was analyzed to extract disease descriptions, formulations, and *materia medica*. Diseases were matched with modern biomedical nomenclature using authoritative references and online medical databases. Plant ingredients were identified and validated using Malaysia Biodiversity Information System (MyBIS), National Parks Board (NParks), and standard ethnobotanical references to confirm scientific names.

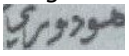
Scientific Validation

To evaluate scientific relevance, each identified *materia medica* was cross-referenced with pharmacological studies indexed in PubMed and Google Scholar. Boolean operators were applied in the search strategy, e.g., ("plant name" OR "scientific name") AND ("in vitro" OR "in vivo" OR "animal study" OR "clinical trial") NOT "meta-analysis." Inclusion criteria encompassed original research articles reporting *in vitro, ex vivo, in vivo*, or clinical studies. Review articles and meta-analyses were excluded before the retrieved studies were screened by title and abstract, and relevant full texts were reviewed. Data were managed in Mendeley for organisation.

RESULTS

Transliteration of text

[1796.19.01] *Sebagai lagi, jika meroyan itu sakit kulit perutnya, ambil daun pekan hitam, tarak (lama) masak jintan, maka disemburkan, maka yang diminum itu daun tambun tahi, ujung melukut, mata kunyit, dan jikalau tiada boleh minum akan buat lempeng beri makan.*

[1796.20.01] *Sebagai lagi ubat menyacakkan peranakan, maka ambil daun*  *segenggam herat (erat), ujung melukut, mata kunyit dan daun hinai (inai) segenggam herat (erat), maka tumbuk beri minum.*

[1796.21.01] *Sebagai lagi jikalau rengat peranakan, maka ambil daun deredap (dedap) dan bonglai, maka bakar tembikar ke api-api, masukkan ke dalam itu beri minum.*

[1796.22.01] *Sebagai lagi jika rengat ari-arinya, maka ambil daun berami, tampal pada ari-arinya.*

[1796.22.02] *(missing word) ambil daun cemperdik (cemperai), giling dengan hujung melukut (ujung melukut), mata kunyit, ramas beri minum maka hampasnya nyiur dibarutkan.*

[1796.22.03] *Sebagai lagi, ambil daun kerak nasi, maka tumbuk dengan hujung melukut (ujung melukut), mata kunyit, ramas beri minum.*

[1796.22.04] *Bab ini sakit ari-ari, maka ambil jerangau, maka pipis lumat-lumat maka tampalkan pada ari-arinya 'afiat.*

[1796.23.01] *Sebagai lagi jikalau meroyan itu berkancing giginya tiada khabar akan dirinya, maka ambil daun leban, tumbuk beri minum, masak jintan hitam, bawang putih kulitnya, semburkan pada tempat sakit itu, masak jintan hitam, bawang putih, dan lada sulah tiga biji, ambil daun ribu-ribu, seni giling lumat-lumat masak ujung melukut, mata kunyit tampalkan di betis 'afiat.*

[1796.24.01] *Sebagai lagi ubat meroyan, salung daunnya, dan daun telinga kerbau, dan kulit limau purut kikis di pohonnya, tumbuk ambil airnya beri minum dan hampasnya buat barut.*

[1796.24.02] *Sebagai lagi ubat meroyan, maka ambil kulit bangkal dan kulit deredap (dedap) sedikit, tumbuk ambil airnya, masak garam, jintan sedikit beri minum maka apabila helat (elat) bukan meroyan, jikalau panas meroyanlah dimakan lagi ubatnya.*

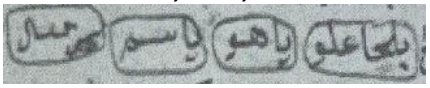
[1796.24.03] *Sebagai lagi ubat meroyan hinggap namanya dimana ia hendak sakit, maka ambil daun benalu, seni disembur pada tempat sakit itu.*

[1796.24.04] *Sebagai lagi meroyan, jikalau melontar*  *ambil daun kayu sugi, pucuk merah dan ambil rumput tinggi panjang*  *dia buah sena maka salit (calit) dibawahnya*.

[1796.26.01] *Bab ini sapulah budak didalam perut sekalipun keluar ulahnya ambil biji labu … (missing word) maka pipis sapukan pada sekalian tubuhnya dan (missing word) ampuan itu nescaya keluar ulahnya.*

[1796.27.01] *Sebagai lagi … (missing word) uri tembuni (uri terselindung), yang maka ambil daun terung asam sehelai yang kukur merah, maka pipiskan airnya air habu (abu), maka suruh minum 'afiat.*

[1796.41.01] *Ini ubat perempuan tiada beranak, maka disurah pada mengkawah putih maka basuh beri jernih airnya tujuh hari perbuat demikian itu jugak, nescaya beroleh anaknya insyaAllah taala inilah rajahnya*



Content Analysis and Materia Medica mentioned in the manuscript

In this study, 15 remedies were identified as related to women's health, of which 14 were plant-based and one

involved ritual practice. These remedies target conditions such as 'meroyan' (postpartum disorders), menyacakkan peranakan (uterine prolapse), and 'rengat *or* sakit ari-arinya' (abdominal pain), underscoring the manuscript's strong emphasis on maternal and reproductive health.

The content analysis revealed the use of several overlapping ingredients across different women's conditions, suggesting a shared pharmacological basis especially for reproductive and postpartum care. Further examination of the medical terminology and interpretation of the diseases enhances our understanding of these overlaps and their therapeutic significance. For instance, certain plants appeared repeatedly in remedies for both postpartum recovery and uterine prolapse, highlighting their perceived importance in Malay traditional medicine.

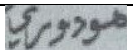
The materia medica recorded in the manuscript consists of 25 distinct ingredients. As categorised, 23 are plant-based and 2 are mineral-based ingredients. Among the identified materia medica, only 9 ingredients such as *bonglai, jerangau, jintan, kayu sugi, kerak nasi, labu, mata kunyit, telinga kerbau and ujung melukut/jerami* were found to be therapeutic in treating women's illnesses based on current scientific research. The materia medica identified through the transliteration text are presented in Tables 1 to 14 with their written names, scientific names, plant parts used, categories, and scientific references relevant to the associated treatments.

The ingredients for 'meroyan (sakit kulit perut)' is tabulated in Table 1 while the ingredients for 'menyacakkan peranakan' is tabulated in Table 2.

Table 1: Ingredients for 1796.19.01 'meroyan (sakit kulit perut)'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Pekan hitam*	Leaf	*Carya illinoinensis*	Plant-based	-
2.	*Jintan*	Seed	*Trachyspermum roxburghianum*	Plant-based	Significant antinociceptive effects (Kalaskar et al., 2024)
3.	*Tambun tahi*	Leaf	*Baccaurea parviflora*	Plant-based	-
4.	*Ujung melukut*	Germinating tip	*Oryza sativa*	Plant-based	Strong antioxidant and anti-inflammatory activities (Thepthanee et al., 2021)
5.	*Mata kunyit*	Germinating tip	*Curcuma longa*	Plant-based	Improve atopic dermatitis-like symptom such as skin inflammation (Jayasinghe et al., 2025)

Table 2: Ingredients for 1796.20.01 'menyacakkan peranakan'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.		Leaf	-	Plant-based	-
2.	*Ujung melukut*	Germinating tip	*Oryza sativa*	Plant-based	Strong antioxidant and anti-inflammatory activities (Thepthanee et al., 2021)
3.	*Mata kunyit*	Germinating tip	*Curcuma longa*	Plant-based	Antioxidant effects (Guerrero - Romero et al., 2020) Bisacurone can modulate inflammation (He at al., 2023)
4.	*Inai*	Leaf	*Lawsonia inermis*	Plant-based	-

The ingredients for ''rengat peranakan' is tabulated in Table 3 while the ingredients for 'rengat atau sakit ari-arinya' is tabulated in Table 4, 5, 6 and 7.

Table 3: Ingredients for 1796.21.01 'rengat peranakan'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Dedap*	Leaf	*Erythrina subumbrans*	Plant-based	-
2.	*Bonglai*	*n.m.*	Zingiber cassumunar	Plant-based	Anti-inflammatory properties of the methanol extract (Khemawoot et al., 2016)

Table 4: Ingredients for 1796.22.01 'rengat atau sakit ari-arinya'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Jerami*	Leaf	*Oryza sativa*	Plant-based	-

Table 5: Ingredients for 1796.22.02 'rengat atau sakit ari-arinya'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Cemperai*	Leaf	*Champereia manillana*	Plant-based	-
2.	*Ujung melukut*	Germinating tip	*Oryza sativa*	Plant-based	-
3.	*Mata kunyit*	Germinating tip	*Curcuma longa*	Plant-based	Reduce the pain in dysmenorrhea (Utami et al., 2020)
4.	*Nyiur*	Dregs	*Cocos nucifera*	Plant-based	-

Table 6: Ingredients for 1796.22.03 for 'rengat atau sakit ari-arinya'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Kerak nasi*	Leaf	*Lindernia crustacea*	Plant-based	Considerable free radical scavenging activity and, antiproliferative and apoptosis-inducing ability (Swargiany et al., 2021)
2.	*Ujung melukut*	Germinating tip	*Oryza sativa*	Plant-based	-
3.	*Mata kunyit*	Germinating tip	*Curcuma longa*	Plant-based	Reduce the pain in dysmenorrhea (Utami et al., 2020)

Table 7: Ingredients for 1796.22.04 'rengat atau sakit ari-arinya'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Jerangau*	*n.m.*	*Acorus calamus*	Plant-based	Antispasmodic activity (Gilani et al., 2006) Anticholinesterase effect in rectus abdominis muscle (Vijayapandi et al., 2012)

The ingredients for 'meroyan (kancing gigi)' is tabulated in Table 8 while the ingredients for 'meroyan' is tabulated in Table 9, 10, 11 and 12.

Table 8: Ingredients for 1796.23.01 'meroyan (kancing gigi)'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Leban*	Leaf	*Vitex pinnata*	Plant-based	-
2.	*Jintan hitam*	Seed	*Nigella sativa*	Plant-based	-
3.	*Bawang putih*	Onion skin	*Allium sativum*	Plant-based	Antiepileptic activity of ethanolic extract (Savairam et al., 2024)
4.	*Lada sulah*	Leaf	*Piper nigrum*	Plant-based	Analgesic and anticonvulsant effects (Bukhari et al., 2013)
5.	*Ribu-ribu*	Leaf	*Lygodium scandens*	Plant-based	-
6.	*Ujung melukut*	Germinating tip	*Oryza sativa*	Plant-based	Prevent lipopolysaccharide-induced brain inflammation and cognitive impairment (Mastinu et al., 2019)
7.	*Mata kunyit*	Germinating tip	*Curcuma longa*	Plant-based	Potential anti-seizure activity (Choo et al., 2021)

Table 9: Ingredients for 1796.24.01 'meroyan'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Salung*	Leaf	*Psychotria viridiflora*	Plant-based	-
2.	*Telinga kerbau*	Leaf	*Blumea balsamifera*	Plant-based	Borneol use in aromatherapy (Ma et al., 2021; Wang et al., 2024)
3.	*Limau purut*	Fruit skin	*Citrus hystrix*	Plant-based	-

Table 10: Ingredients for 1796.24.02 'meroyan'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Bangkal*	Fruit skin	*Ludekia bernardoi*	Plant-based	-
2.	*Dedap*	Fruit skin	*Erythrina subumbrans*	Plant-based	-
3.	*Garam*	-	Sodium chloride	Mineral-based	-
4.	*Jintan*	Seed	*Trachyspermum roxburghianum*	Plant-based	Antioxidant effect (Kalaskar et al., 2024)

Table 11: Ingredients for 1796.24.03 for 'meroyan'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Benalu*	Leaf	*Hemigraphis colorata*	Plant-based	-

Table 12: Ingredients for 1796.24.04 'meroyan'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Kayu sugi*	Leaf	*Salvadora persica*	Plant-based	Antidepressant-like effects when combined with date palm (Youssef et al., 2022)
2.	*Pucuk merah*	*n.m.*	*Syzygium oleana*	Plant-based	-
3.	*Rumput*	*n.m.*	Unspecified	Plant-based	-
4.	*Sena*	Fruit	*Pterocarpus indicus*	Plant-based	-

The ingredients for 'ulah budak di dalam perut' is tabulated in Table 13 The ingredients for 'uri tembuni' is tabulated in Table 14.

Table 13: Ingredients for 1796.26.01 'ulah budak di dalam perut'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Labu*	Seed	*Lagenaria siceraria*	Plant-based	Immune-enhancing adjuvant activity (Wusiman et al., 2022)

Table 14: Ingredients for 1796.27.01 'uri tembuni'

No.	Materia medica	Plant part	Scientific name	Category	Scientific References
1.	*Terung asam*	Leaf	*Solanum lasiocarpum*	Plant-based	-
2.	*Air abu*	-	Charcoal	Mineral-based	-

DISCUSSION

The analysis of Malay medical manuscript MSS 1796 revelead nine illnesses related to women's physical and non physical illness. The diseases are 'meroyan (sakit kulit perut)', 'menyacakkan peranakan', 'rengat peranakan', 'rengat atau sakit ari-arinya', 'meroyan'(kancing gigi), 'meroyan', 'ulah budak di dalam perut', 'uri tembuni' and, 'perempuan tiada beranak'. To contextualise the remedies documented in *MSS 1796*, it is essential to elaborate on the Malay disease terms and their corresponding medical interpretations. The manuscript references several conditions related to postpartum and reproductive health, each carrying specific ethnomedical meanings. For example, 'meroyan (sakit kulit perut)' refers to postpartum hives, characterised by itchy, red rashes that may appear after childbirth, even in women without a prior history of allergies, often due to hormonal changes (Thourani, 2022).

The term 'menyacakkan peranakan' describes efforts to reposition the uterus, aligning with the modern understanding of uterine prolapse, which is a condition in which weakened pelvic muscles and ligaments allow the uterus to descend into the vaginal canal, with childbirth being a major risk factor (Johns Hopkins Medicine, n.d.). Similarly, 'rengat peranakan' denotes postpartum cramping associated with uterine involution. These contractions resemble menstrual cramps and are commonly managed using analgesics and non-steroidal anti-inflammatory drugs (NSAIDs) (Deussen et al., 2020). In the manuscript, remedies featuring herbs such as *dedap* and *bonglai* are prescribed for this condition. The phrase 'rengat atau sakit ari-arinya' refers more broadly to lower abdominal pain, which corresponds to dysmenorrhea that is a menstrual cramp that typically affect the lower abdomen and may radiate to the back or thighs (Nagy et al., 2023).

Interestingly, the manuscript also employs 'meroyan' to describe postpartum depression, a serious mental health condition characterised by anxiety and depressive symptoms following childbirth (OASH, 2023). Another term, 'ulah budak di dalam perut', relates to pregnancy-associated complications that may endanger the mother, the baby, or both. Meanwhile, 'uri' and 'tembuni' both denote the placenta, reflecting the manuscript's careful attention to maternal physiology (PRPM, 2017a; PRPM, 2017b).

The manuscript further addresses infertility through the phrase 'perempuan tiada beranak', indicating a woman's inability to conceive. Notably, this prescription departs from plant-based treatments and incorporates spiritual healing: the recitation of Qur'anic verses (*surah*) into a *mengkawah putih* (a white pot), with the resulting water consumed or used over seven days. This reflects the integration of spiritual belief and therapeutic practice within Malay ethnomedicine, particularly in matters concerning fertility.

Together, these terms highlight the manuscript's sophisticated ethnomedical vocabulary and provide an essential interpretive foundation for analysing the corresponding remedies. Understanding the cultural and medical meanings embedded in these Malay terms allows for a more accurate interpretation of *MSS 1796*, thereby strengthening the discussion on how traditional formulations relate to contemporary biomedical concepts in women's health. The diseases with the highest number of remedies are 'rengat atau sakit ari-arinya' and 'meroyan', with four remedies each. Furthermore, the presence of a spiritual remedy highlights the integration of spiritual healing alongside physical treatments, which is the characteristic of traditional medical systems.

Among the most prominent ingredients recorded in the manuscript is *Oryza sativa*, often referred to as *ujung melukut* (tip of the rice husk) or *jerami* (rice straw). Despite being agricultural byproducts, these parts of the rice plant are incorporated into women's health formulations, reflecting a resourceful and holistic approach to healing. Their repeated use in conditions such as 'meroyan' *(sakit kulit perut)*, 'menyacakkan peranakan', and 'rengat atau sakit ari-arinya' underscores their therapeutic importance. Modern pharmacological studies have confirmed that *Oryza sativa* possesses significant anti-inflammatory and antioxidant properties (Thepthanee et al., 2021), suggesting scientific plausibility behind its traditional application.

Closely following is *Curcuma longa* (*mata kunyit*), which appears in four formulations for similar conditions. Turmeric is widely documented for its anti-inflammatory and antioxidant properties, with modern evidence supporting its efficacy in alleviating dysmenorrhea and reducing postpartum pain (Guerrero-Romero et al., 2020; Utami et al., 2020; He et al., 2023; Jayasinghe et al., 2025). Its role in Malay traditional medicine resonates with its use in Ayurveda and Traditional Chinese Medicine (TCM), where turmeric is similarly employed in managing reproductive and postpartum conditions.

Another notable ingredient is *Trachyspermum roxburghianum* (*jintan*), recorded for formulations addressing *meroyan (sakit kulit perut)*. Research has demonstrated its antinociceptive and antioxidant potential (Kalaskar et al., 2024), which may account for its traditional application in alleviating abdominal discomfort after childbirth. Meanwhile, other ingredients such as *Zingiber cassumunar (bonglai)*, *Acorus calamus (Jerangau)*, Salvadora *persica (kayu sugi)*, *Lindernia*

crustacea (kerak nasi), *Lagenaria siceraria (labu)*, and *Blumea balsamifera (telinga kerbau)* appear less frequently but nonetheless carry significant ethnopharmacological value, warranting further study.

The prominence of remedies targeting conditions such as *meroyan* and *rengat peranakan* underscores the strong cultural emphasis on postpartum health in Malay society. The period of confinement ('berpantang') is viewed as crucial to restoring women's bodily balance and preventing long-term illness (Teh et al., 2021). Herbs and plant-based ingredients serve not only as physical remedies but also as symbolic protectors of the female body, believed to realign internal balance, restore vitality, and safeguard fertility (Akbaribazm et al., 2021; Fareez, 2024).

This focus on postpartum care aligns with practices in other traditional medical systems. In Traditional Chinese Medicine (TCM), the postpartum period is considered a vulnerable "cold" state requiring "warming" foods and herbs, while in Ayurveda, confinement is treated with herbs to restore the uterus, aid lactation, and rejuvenate energy (Ho et al., 2011; Akbaribazm et al., 2021; Patibandla et al., 2024). Thus, Malay formulations share a global ethnomedical recognition of postpartum care as a cornerstone of women's health.

The findings from MSS 1796 have implications beyond cultural preservation. Further studies into the formulations targeting the conditions could inspire the development of complementary therapies or nutraceuticals for women's health, particularly in maternal care where integrative approaches are increasingly valued (Lunze et al., 2015; Schürger et al., 2018). Additionally, given the persistence of maternal health disparities in many regions, these formulations could offer accessible and culturally acceptable alternatives for communities with limited access to conventional medicine (Lunze et al., 2015).

CONCLUSION

This study revealed the significance of Malay medical manuscript MSS 1796 in preserving traditional knowledge on women's health documenting eight conditions and fourteen plant based prescriptions. Plants such as *ujung melukut* (*Oryza sativa)*, *mata kunyit* (*Curcuma longa)*, and *Jintan* (*Trachyspermum roxburghianum)* illustrate both cultural relevance and pharmacological potential, particularly in management of postpartum disorders, uterine prolapse and abdominal pain. The findings highlights the central importance of women's health in Malay medical tradition while revealing opportunities for further scientific validation. By bridging the traditional ethnomedical knowledge with contemporary biomedical insights, this study contributes to the preservation of Malay heritage and the development of integrative approaches to women's health.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Akbaribazm, M., Goodarzi, N., & Rahimi, M. (2021). Female infertility and herbal medicine: An overview of the new findings. *Food Science & Nutrition*, *9*(10), 5869–5882. https://doi.org/10.1002/fsn3.2523

Deussen, A. R., Ashwood, P., Martis, R., Stewart, F., & Grzeskowiak, L. E. (2020). Relief of pain due to uterine cramping/involution after birth. *Cochrane Database of Systematic Reviews*, *2020*(10). https://doi.org/10.1002/14651858.cd004908.pub3

Fareez, A. (2024, March 18). *Postpartum Nutrition in Malaysia - LifeCare*. LifeCare. https://lifecare.com.my/ukiyo-article/postpartum-nutrition-in-malaysia/

Geller, A., Salganicoff, A., & Burke, S. P. (2025, February 14). *Overview of Selected Women's Health Conditions*. Nih.gov; National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK612403/

Gilani, A. U. H., Shah, A. J., Ahmad, M., & Shaheen, F. (2006). Antispasmodic effect of Acorus calamus Linn. is mediated through calcium channel blockade. *Phytotherapy Research*, *20*(12), 1080–1084. https://doi.org/10.1002/ptr.2000

Grant, A. (2023, August 4). *FDA Approves First Oral Treatment for Postpartum Depression*. FDA. https://www.fda.gov/news-events/press-announcements/fda- approves-first-oral-treatment-postpartum-depression

Guerrero-Romero, F., Simental-Mendía, L. E., Martínez-Aguilar, G., Sánchez-Meraz, M. A., & Gamboa-Gómez, C. I. (2020). Hypoglycemic and antioxidant effects of five commercial turmeric (Curcuma longa) supplements. *Journal of Food Biochemistry*, *44*(9). https://doi.org/10.1111/jfbc.13389

Hadi Solehim. (2024, March 25). *National Library programmes aim to analyse 5,252 Malay manuscripts*. New Straits Times; New Straits Times Press (M) Bhd. https://www.nst.com.my/news/nation/2024/03/1030385/national-library- programmes-aim-analyse-5252-malay-manuscripts

Harun Mat Piah. (2006). *Traditional Malay Medicine from Old Malay Texts*. International Journal of the Malay World and Civilisation (Iman) 3(3), 2015: 3 - 17. https://www.ukm.my/jatma/wpcontent/uploads/makalah/IMAN%203(3)%20Sept%202015%20-%20%5B01%5D%20harun.pdf

He, C., Miyazawa, T., Abe, C., Ueno, T., Suzuki, M., Mizukami, M., Kurihara, K., & Toda, M. (2023). Hypolipidemic and Anti-Inflammatory Effects of Curcuma longa- Derived Bisacurone in High-Fat Diet-Fed Mice. *International Journal of Molecular Sciences*, *24*(11). https://doi.org/10.3390/ijms24119366

Ho, M., Li, T.-C., & Su, S.-Y. (2011). The Association between Traditional Chinese Dietary and Herbal Therapies and Uterine Involution in Postpartum Women. *Evidence-Based Complementary and Alternative Medicine*, *2011*, 1–9. https://doi.org/10.1155/2011/918291

Höller, M., Steindl, H., Dimitri Abramov-Sommariva, Kleemann, J., Alexey Loleit, Abels, C., & Stute, P. (2024). Use of Vitex agnus-castus in patients with menstrual cycle disorders: a single-center retrospective longitudinal cohort study. *Archives of Gynecology and Obstetrics*, *309*, 2089–2098. https://doi.org/10.1007/s00404-023-07363-4

Jayasinghe, A. M. K., Kirindage, K. G. I. S., Kim, S.-H., Lee, S., Jung, K., Shim, S.-Y., & Ahn, G. (2025). Protective effect of Curcuma longa L. leaves and pseudostems extract against 1-chloro-2,4-dinitrobenzene-induced atopic dermatitis in BALB/c mice. *Journal of Ethnopharmacology*, *338*(Pt 3), 119138. https://doi.org/10.1016/j.jep.2024.119138

Intan Shahhira Noorain, & Mardhiah Mohammad. (2021). *The Medical Use of Germinating Parts of Plants; Mata Kunyit, Sunti Halia and Lemukut in Malay Medical Manuscripts*. International Journal of Allied Health Sciences, 5(5), 2396–2406. https://doi.org/10.31436/ijahs.v5i5.743

Jabatan Perhutanan Semenanjung Malaysia. (n.d.). *Sembong*. Jabatan Perhutanan Semenanjung Malaysia. https://www.forestry.gov.my/images/info-perhutanan/tumbuhan-ubatan/Sembong_en.pdf

Johns Hopkins Medicine. (n.d.). *Uterine Prolapse*. Www.hopkinsmedicine.org. https://www.hopkinsmedicine.org/health/conditions-and-diseases/uterine-prolapse

Kalaskar, M., Gavit, A., Prabhu, S., Gagarani, M., Ugale, V., Khadse, S., Ayyanar, M., Surana, S., Tatiya, A., & Gurav, S. (2024). Chemical composition, antioxidant, antimicrobial, and wound healing effects of Trachyspermum roxburghianum (DC.) H. Wolff essential oil: An in vivo and in silico approach. *Journal of Ethnopharmacology*, *327*. https://doi.org/10.1016/j.jep.2024.118055

Khemawoot, P., Hunsakunachai, N., Anukunwithaya, T., Bangphumi, K., Ongpipattanakul, B., Jiratchariyakul, W., Soawakontha, R., Inthachart, T., Ayudhya, N., Koontongkaew, S., & Poachanukoon, O. (2016). Pharmacokinetics of Compound D, the Major Bioactive Component of Zingiber cassumunar, in Rats. *Planta Medica*, *82*(13), 1186–1191. https://doi.org/10.1055/s-0042-104658

Kumar, S., Dhir, A., Talwar, S., Chakraborty, D., & Kaur, P. (2021). What drives brand love for natural products? The moderating role of household size. *Journal of Retailing and Consumer Services*, *58*(58), 102329. https://www.sciencedirect.com/science/article/pii/S0969698920313370

Luan, H., Liu, Q., Guo, Y., Fan, H., A., S., & Lin, J. (2025). Effects of soy isoflavones on menopausal symptoms in perimenopausal women: A systematic review and meta-analysis. *PeerJ*, *13*, e19715. https://doi.org/10.7717/peerj.19715

Lunze, K., Higgins-Steele, A., Simen-Kapeu, A., Vesel, L., Kim, J., & Dickson, K. (2015). Innovative approaches for improving maternal and newborn health - A landscape analysis. *BMC Pregnancy and Childbirth*, *15*(1). https://doi.org/10.1186/s12884-015-0784-9

Ma, R., Su, P., Ma, Q., Guo, J., Chen, S., Jin, B., Zhang, H., Tang, J., Zhou, T., Xiao, C., Cui, G., & Huang, L. (2021). Identification of (-)-bornyl diphosphate synthase from Blumea balsamifera and its application for (-)-borneol biosynthesis in Saccharomyces cerevisiae. *Synthetic and Systems Biotechnology*, *7*(1), 490–497. https://doi.org/10.1016/j.synbio.2021.12.004

Meresman, G. F., Götte, M., & Laschke, M. W. (2021). Plants as source of new therapies for endometriosis: a review of preclinical and clinical studies. *Human reproduction update*, *27*(2), 367–392. https://doi.org/10.1093/humupd/dmaa039

Mohd Affendi Mohd Shafri. (2021). Sakti Index: Towards Scientific Evaluation of Malay Medical Manuscripts. *International Journal of Allied Health Sciences*, *5*(5), 2351–2363. https://journals.iium.edu.my/ijahs/index.php/IJAHS/article/view/740/632

Nagy, H., Carlson, K., & Khan, M. A. (2023, November 12). *Dysmenorrhea*. PubMed; StatPearls Publishing.

https://www.ncbi.nlm.nih.gov/books/NBK560834/

OASH. (2023, October 17). *Postpartum Depression | Office on Women's Health*. Womenshealth.gov. https://womenshealth.gov/mental-health/mental-health- conditions/postpartum-depression

Patibandla, S., Gallagher, J. J., Patibandla, L., Ansari, A. Z., Qazi, S., & Brown, S. F. (2024). Ayurvedic Herbal Medicines: A Literature Review of Their Applications in Female Reproductive Health. *Cureus*, *16*(2). https://doi.org/10.7759/cureus.55240

Perpustakaan Negara Malaysia. (2024). *e-Bibliografi Beranotasi Manuskrip Melayu - KitabTib Khazanah Kebitaraan Warisan Bangsa*. Www.pnm.gov.my.https://www.pnm.gov.my/pnm/resources/pdf%20file/Latest%20e-Bibliografi%20Beranotasi%20Manuskrip%20Melayu%20-%20Kitab%20Tib- compressed.pdf

Pusat Rujukan Persuratan Melayu (PRPM). (2017a). *Carian Umum*. Dbp.gov.my. https://prpm.dbp.gov.my/cari1?keyword=uri

Pusat Rujukan Persuratan Melayu (PRPM). (2017b). *Carian Umum*. Dbp.gov.my. https://prpm.dbp.gov.my/cari1?keyword=tembuni

Schürger, N., Klein, E., Hapfelmeier, A., Kiechle, M., & Paepke, D. (2018). Demand for integrative medicine among women in pregnancy and childbed: a German survey on patients' needs. *BMC Complementary and Alternative Medicine*, *18*(1). https://doi.org/10.1186/s12906-018-2249-y

Su S. Y. (2022). Editorial: Herbal medicines in women's lives. *Frontiers in pharmacology*, *13*, 1003241. https://doi.org/10.3389/fphar.2022.1003241

Swargiary, A., Verma, A. K., Singh, S., Roy, M. K., & Daimari, M. (2020). Antioxidant and Antiproliferative Activity of Selected Medicinal Plants of Lower Assam, India: An In Vitro and In Silico Study. *Anti-Cancer Agents in Medicinal Chemistry*, *21*(2), 267–277. https://doi.org/10.2174/1871520620666200719000449

Takahashi, T. A., & Johnson, K. M. (2015). Menopause. *Medical Clinic of North America*, *99*, 521–534. https://doi.org/10.1016/j.mcna.2015.01.006

Teh, K., Quek, I. P., & Tang, W. E. (2021). Postpartum dietary and physical activity-related beliefs and behaviors among women with recent gestational diabetes mellitus: a qualitative study from Singapore. *BMC Pregnancy and Childbirth*, *21*(1). https://doi.org/10.1186/s12884-021-04089-6

Thepthanee, C., Liu, C.-C., Yu, H.-S., Huang, H.-S., Yen, C.-H., Li, Y.-H., Lee, M.-R., & Liaw, E.-T. (2021). Evaluation of Phytochemical Contents and In Vitro Antioxidant, Anti-Inflammatory, and Anticancer Activities of Black Rice Leaf (Oryza sativa L.) Extract and Its Fractions. *Foods (Basel, Switzerland)*, *10*(12). https://doi.org/10.3390/foods10122987

Thourani, A. (2022, October 10). *Postpartum Hives: Symptoms, Causes and Treatments That Work*. Revival Research Institute, LLC. https://revivalresearch.org/blogs/postpartum-hives/

Utami, R. B., Damayanti, D. F., & Rodiah, D. (2020). The effectiveness of curcuma longa drink in decreasing the intensity of dysmenorrhea. *Biomedical and Pharmacology Journal*, *13*(4), 2055–2060. https://doi.org/10.13005/BPJ/2085

Vallée, A., & Lecarpentier, Y. (2020). Curcumin and Endometriosis. *International Journal of Molecular Sciences*, *21*(7). https://doi.org/10.3390/ijms21072440

Vijayapandi, P., Annabathina, V. K., SivaNagaSrikanth, B., Manjunath, V., Boggavarapu, P., Mohammed P, A. K., RajendraPrasad, K., & Kumarappan, C. T. (2012). In vitro anticholinergic and antihistaminic activities of Acorus calamus Linn. leaves extracts. *African Journal of Traditional, Complementary, and Alternative Medicines : AJTCAM / African Networks on Ethnomedicines*, *10*(1), 95–101. https://doi.org/10.4314/ajtcam.v10i1.13

Wan Ali W. M. (2017). *Malay Ingenuity in the Use of Local Materials for the Production of Malay Manuscripts*. Perpustakaan Negara Malaysia. https://www.pnm.gov.my/pnm/resources/pdf%20file/PKMM/Malay%20Ingenuity%20n%20the%20use%20of%20localmaterials%20for%20the%20production%20 of%20malay%20manuscripts.pdf

Wang, K., Hu, X., Xie, X. L., Huang, M., Wang, D., & Yu, F. L. (2024). Phytocosmetic potential of Blumea balsamifera oil in mitigating UV-induced photoaging: Evidence from cellular and mouse models. *Journal of Ethnopharmacology*, *334*. https://doi.org/10.1016/j.jep.2024.118535

Wusiman, A., Rexiati, S., Aziz, M., Cheng, X., Mai, Z., Abulaiti, A., Wutikuer, A., Rozi, P., Abuduwaili, A., & Abula, S. (2022). Preparation and sulfate modified of Lagenaria siceraria (Molina) Standl polysaccharide and its immune-enhancing adjuvant activity. *Poultry*

Science,*101*(11). https://doi.org/10.1016/j.psj.2022.102112

Youssef, B., Ramadan, K. S., ElShebiney, S., & Ibrahim, E. A. (2022). Antidepressant-like effects of aqueous extracts of miswak (Salvadora persica) and date palm (Phoenix dactylifera) on depression-like behaviors using CUMS model in male rats. *Journal of Food Biochemistry*, *46*(8). https://doi.org/10.1111/jfbc.14164

Zhou, X., Afzal, S., Wohlmuth, H., Münch, G., Leach, D., Low, M., & Li, C. G. (2022). Synergistic Anti-Inflammatory Activity of Ginger and Turmeric Extracts in Inhibiting Lipopolysaccharide and Interferon-γ-Induced Proinflammatory Mediators. *Molecules*, *27*(12), 3877. https://doi.org/10.3390/molecules27123877

REVIEW ARTICLE

Time-Lapse Imaging (TLI) versus Preimplantation Genetic Testing (PGT): A Scoping Review on Embryo Assessment Strategies

Azantee Yazmie Abdul Wahab[1,*], Muhammad Izzat Zhulkefli[1]

[1]Department of Biomedical Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
Time-lapse imaging; preimplantation genetic testing; embryo assessment; clinical outcomes

ABSTRACT

Background: Embryo selection remains a critical determinant in the success of assisted reproductive technologies (ART). While conventional morphological assessment is widely used, it lacks predictive precision and is prone to subjectivity. Two advanced technologies; Time-Lapse Imaging (TLI) and Preimplantation Genetic Testing (PGT) have emerged to enhance embryo viability assessment, yet their comparative effectiveness remains uncertain. This scoping review aimed to evaluate and compare the effectiveness of TLI and PGT in embryo selection for transfer, their impact on clinical outcomes, and the strengths and limitations of each method. **Methods:** A scoping review methodology based on PRISMA for scoping review 2020 guidelines was employed. Four databases PubMed, Scopus, Springer Link, and Cochrane Library were searched for articles published between 2014 and 2024. **Results:** A total of 20 eligible studies were included and analysed. TLI studies reported improved morphological assessment and non-invasive, real-time observation, with pregnancy rates ranging between 27.8% to 68% and live birth rates up to 45.9%. PGT, particularly PGT-A, showed superior outcomes in terms of implantation up to 76.8% and live birth rates up to 78.3%, especially among patients with recurrent pregnancy loss or advanced maternal age. However, PGT carries the risk of embryo damage due to its invasive nature and is associated with higher costs and longer processing times. TLI offers a safer and cost effective option for embryo selection with minimal risk to embryo integrity. TLI's integration with software tools such as iDA scores has further enhanced its diagnostic capability. Nonetheless, in specific clinical scenarios like recurrent miscarriages or known genetic disorders, PGT remains indispensable. **Conclusion:** Both TLI and PGT offer significant advantages in embryo assessment. While TLI is favorable for routine, non-invasive monitoring, PGT provides critical genetic insights.

INTRODUCTION

Assisted reproductive technologies (ART) have transformed fertility treatment, but the success rate remains challenging, with a 2016 global success rate of 30.9% (Kupka et al., 2024). The conventional method of embryo selection, based on subjective morphological assessments, has limitations in accurately predicting a successful pregnancy. This has led to the development of new technologies, notably time-lapse imaging (TLI) and preimplantation genetic testing (PGT), which aim to improve the identification of viable embryos. While both TLI and PGT are designed to increase implantation rates and reduce miscarriages, they employ different mechanisms, and the debate over their comparative effectiveness is ongoing.

TLI is a non-invasive technology that allows embryologists to monitor embryo development continuously in a stable environment. By capturing images at regular intervals, TLI creates a video record of an embryo's growth, providing detailed insights into its developmental kinetics and morphology (Wang et al., 2023). This method enhances the traditional scoring system, as it allows for a more comprehensive and objective assessment of key parameters like cleavage timing and cell division patterns. Advanced algorithms, such as the Intelligent Data Analysis (IDA) score, further improve TLI's predictive accuracy by using AI to analyze large datasets and reduce human bias (Papamentzelopoulou et al., 2024). However, TLI's limitations include its dependence on high-quality data and a lack of consideration for patient-specific factors or genetic health.

PGT is an invasive procedure that involves screening embryos for genetic and chromosomal abnormalities before they are transferred to the uterus. PGT helps select the healthiest embryos, increasing the likelihood of a successful pregnancy and reducing the risk of genetic disorders and miscarriages, especially for patients with known hereditary risks (Sravya Gudapati et al., 2024). The accuracy of PGT has significantly improved with advancements from older methods like fluorescence in situ hybridization to modern

* *Corresponding author.*
E-mail address: yazmie@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

techniques such as next-generation sequencing. Despite its benefits in improving outcomes, PGT is an invasive procedure that requires an embryo biopsy, which carries a potential risk to embryo viability (Burks et al., 2021). It also increases the overall cost and complexity of the treatment, making it less accessible for some patients.

The central question in modern ART is which method is more effective in selecting the most viable embryos for transfer. Both TLI and PGT have distinct advantages and limitations. The aim of this review is to provide a comprehensive analysis of the existing literature, synthesizing evidence to evaluate whether one technique demonstrates superiority, or if a combined approach yields the best outcomes. This analysis will offer insights into which embryo selection method is most effective, and whether the invasive nature of PGT is justified by improved results compared to the less invasive TLI. Ultimately, the findings will contribute to the development of more precise, reliable, and patient-centered embryo selection strategies, thereby enhancing success rates.

MATERIALS AND METHODS

Study design

This research utilizes a scoping review design to comprehensively map the existing evidence on the effectiveness of TLI and PGT in assessing embryo viability for assisted reproductive technology. The framework included the development of research objectives aligned with the research problems, the selection of articles that met the review's goals, and a screening process to ensure compliance with the inclusion and exclusion criteria, along with the availability of full-text articles.

Search Method

Relevant research articles were retrieved from four medically and scientifically credible online databases: PubMed, Scopus, Springer Link and Cochrane Library. The following keywords were used to identify relevant publications: 'time-lapse imaging', 'time-lapse monitoring', 'TLI', 'preimplantation genetic testing', 'preimplantation genetic screening', 'PGT', 'embryo assessment', 'in vitro fertilization', 'intracytoplasmic sperm injection', 'invasive', 'embryo selection', 'embryo grading', 'fertilization rate', 'blastocyst rate', 'implantation rate', 'pregnancy rate', 'miscarriage rate', 'live birth rate', 'clinical outcome', 'cost', 'invasiveness', 'time', and 'accuracy'. Additionally, in order to further clarify and broaden the search parameters while using databases, Boolean operators (AND, OR) and asterisks (*) were utilized in combination with the search keywords.

Inclusion and exclusion criteria

Several criteria were established and considered during the selection of relevant papers for this scoping review. Articles that failed to meet the inclusion criteria outlined in Table 1 were excluded from the analysis.

Selection of studies

The research papers were selected following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Reviews (PRISMA-ScR) standards. These standards provide guidelines for choosing, evaluating, and synthesizing studies in scoping reviews, and the most recent statement incorporates improvements in these approaches (Page et al., 2021). Articles published between 2014 and 2024 were extracted from online databases and screened. After eliminating duplicates from various sources, the remaining articles underwent title and abstract screening to filter out irrelevant studies. Subsequently, the full-text versions of the selected papers were retrieved. Next, the information in those articles was thoroughly examined to analyze the data on the effectiveness of TLI or PGT in embryo assessment. During the screening process, articles that did not meet the inclusion criteria or conformed to the exclusion criteria were excluded. The PRISMA extension for scoping reviews is presented in Figure 1.

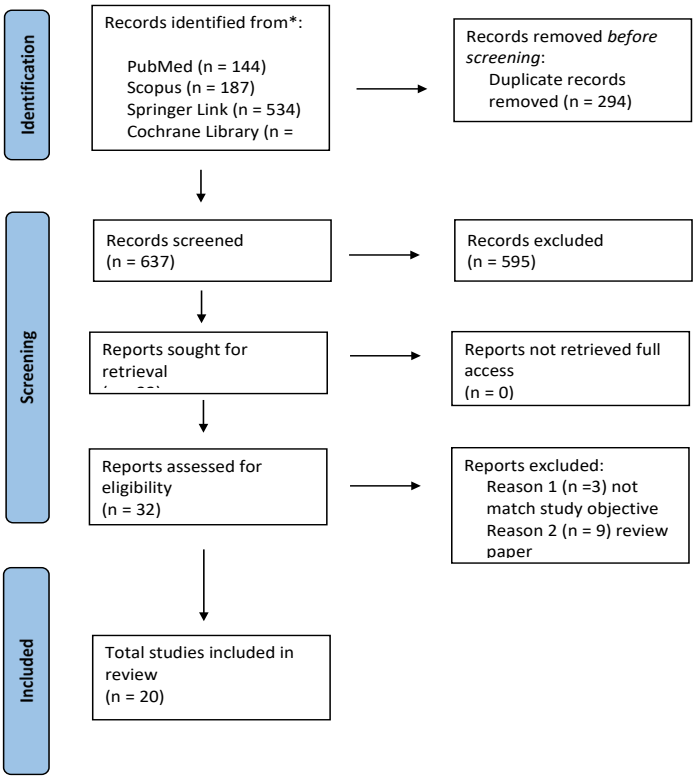


Figure 1: PRISMA flow diagram

Data extraction

The final phase for methodology is data retrieval from the chosen articles. The data were extracted and analysed through meticulous reading and examination to synthesize key insights of the studies according to objectives and identify advantages and limitations of TLI and PGT. The findings from the final selection were extracted presented in Table 2. Here, the characteristics of each study were summarized and were presented with their references, ordered by publication year (latest to oldest). A primary focus across these studies is the evaluation of embryo viability and clinical outcomes, alongside findings of the strengths and limitations of each technology.

RESULTS

A comprehensive search yielded, a total of 931 potential articles were identified from the selected database. After removing duplicates, 294 references remained. A total of 32 articles were eligible based on inclusion and exclusion criteria. Among the eligible articles, only 20 publications match with the objective of this study. There are 10 studies specifically discuss TLI, 9 focus on PGT, and 1 observational study compares the costs of both techniques

Effectiveness TLI and PGT in selecting viable embryos for transfer

Figure 2 presents a comparative analysis of TLI and PGT across three aspects: fertilization rate, blastocyst formation rate, and embryo quality. PGT demonstrated superior performance in fertilization rate and embryo quality, with a mean fertilization rate of 67.52% compared to 64.02% for TLI, and a significantly higher mean embryo quality score of 78.84% versus 65.08% for TLI. Conversely, TLI showed a slightly higher average blastocyst formation rate at 55.2%, compared to 51.46% in the PGT group.

Impact of TLI and PGT on clinical outcomes

Figure 3 presents a comparative analysis of TLI and PGT across four aspects: implantation rate, pregnancy rate, miscarriage rate, and live birth rate. PGT consistently demonstrates superior clinical outcomes in key indicators of ART success. For implantation and pregnancy rates, the mean percentages for PGT are 63.78% and 59.17%, respectively, which are significantly higher than TLI's 33.14% and 47.22%. Furthermore, PGT is associated with a greater likelihood of live birth, with a mean rate of 51.19% compared to 39.6% for TLI. In contrast, the TLI group showed a slightly lower mean miscarriage rate of 8.2% compared to 11.33% for PGT.

Strength, limitations, and potential drawbacks of both TLI and PGT

Table 3 presents a comparative analysis of TLI and PGT in terms of time, invasiveness, accuracy, and cost. TLI is a non-invasive, efficient, and relatively affordable method that provides real-time insights into an embryo's developmental kinetics without disrupting its environment. In contrast, PGT is a highly accurate but invasive procedure that genetically screens embryos, a process that is more time-consuming and significantly more expensive. The invasiveness of PGT carries a risk of harming the embryo, potentially compromising its viability. While TLI relies on morphokinetic parameters to predict success, PGT directly assesses chromosomal health; however, its ability to detect complex genetic abnormalities, such as mosaicism, remains limited due to sampling constraints.

Table 1: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
a) Studies that are published in English.	a) Studies that are published in languages other than English.
b) Studies that are published from year 2014 to 2024.	b) Studies that are published before year 2014.
c) Studies that are related to embryo assessment with TLI or PGT.	c) Studies that are not related to embryo assessment.
d) Full-text articles.	d) Unpublished studies, conference paper, grey literature, letters, editorials, reviews (systematic, scoping, narrative reviews), and meta-analysis.
e) Clinical studies that consists of related specific keywords and related with the objectives.	e) Studies that conducted on animals

Table 2: Findings of TLI and PGT in embryo assessment.

Author	Study Design	Assessment method	Findings		
			Embryo Viability	Clinical outcomes	Strength & limitations
Xu et al., (2024)	Retrospective	PGT-A with NGS	(i)High quality embryo: 67.7% (ii)19 mosaic embryo used for transfer	(i)Pregnancy Rate: 64.2% (ii)Miscarriage Rate: 17.2% (iii)Live Birth Rate: 51.3%	PGT enhanced selection of euploid embryos, and increase the chances of successful pregnancy were enhanced, particularly in patients with a history of recurrent miscarriages
Illingworth et al.(2024)	Randomized Controlled Trial (RCT)	TLI with iDA score	(i)Fertilization rate: 60.95% (ii)Blastocyst Formation Rate: 68.5% (iii)Mean viable embryo iDA score transfer (9.09±0.68)	(i)Pregnancy Rate: 46.5% (ii)Live Birth Rate: 39.8%	TLI took only 21.3 seconds to evaluate embryos
Bhide et al., (2024)	RCT	TLI	Fertilization rate: 60.66%	(i)Pregnancy Rate: 42.2% (ii)Miscarriage Rate: 8.7% (iii)Live Birth Rate: 33.7%	TLI only requires monitoring and recording of the embryos over time, which is less labor intensive and requires fewer human resources.
Harris et al.,(2024)	Retrospective	PGT-A	Patient age under 35 (i)Fertilization rate: 62.94% (ii)Blastocyst rate: 54.03% Patient age 35-37 (i)Fertilization rate: 63.58% (ii)Blastocyst formation rate: 52.73% Patient age 38-40 (i)Fertilization rate: 62.96% (ii)Blastocyst formation rate: 46.08%	Patient age under 35 (i)Pregnancy rate: 54.3% (ii)Miscarriage rate: 19.5% (iii)Live birth rate: 43.3% Patient age 35-37 (i)Pregnancy Rate: 64.8% (ii)Live Birth Rate: 43.0% (iii)Miscarriage Rate: 8% Patient age 38-40 (i)Pregnancy rate: 53.3% (ii)Miscarriage rate: 19.5% (iii)Live birth rate: 43.3%	PGT allows identification and prioritization of chromosomally normal embryos which may lead to better clinical outcomes.
Katz-Jaffe et al.,	Prospective	PGT-A	(i)Fertilization Rate: 59.42%	(i)Implantation Rate: 76.8%	-

(2024)				(ii)Miscarriage Rate: 4% (iii)Live Birth Rate: 78.3%	
Zheng et al., (2023)	Retrospective	PGT-A with NGS	(i)Fertilization rate: 70.75% (ii)High quality embryo: 89.92%	(i)Pregnancy Rate: 63.55% (ii)Live Birth Rate: 55.26%	-
Meng et al., (2022)	RCT	TLI with KID score	Mean top quality embryo: Control (7.15 ± 4.42) vs TLI (5.62 ±3.07)	(i)Pregnancy Rate: 49.18% (ii)Miscarriage Rate: 3.33% (iii)Live Birth Rate: 45.9%	TLI offer a real-time observation and captures detailed developmental and abnormalities such as direct cleavage and reverse cleavage.
Singh et al., (2022)	Retrospective	PGT-A with NGS	Fertilization rate: 69.54%	(i) Implantation Rate: 66.6% (ii) Pregnancy Rate: 81.8% (iii) Miscarriage Rate: 14.8% (iv) Live Birth Rate: 66.6%	(i)Abnormalities detected (Gene mutatiom:40% Chromosal abnormalities; 54.5% Mosaic: 5.45%) (ii)The application of NGS in PGT process provides a comprehensive and accurate screening of all 24 chromosomes
Sui et al., (2020)	RCT	PGT-A with SNP	Fertilization rate: PGT (70%)	(i)Implantation Rate: 48.51% (ii)Pregnancy Rate: 51.3% (iii)Miscarriage Rate: 1.74% (iv)Live Birth Rate: 43.48%	(i)A significant proportion of cycles (50.27%) and patients (24.27%) in the PGT-A group failed to reach embryo transfer due to the loss of euploid embryos during blastocyst biopsy. (ii)PGT-A process, including blastocyst biopsy, genetic screening, embryo vitrification, and counselling, can take a lot of time. This process may slow down embryo transfer and increase stress for patients already facing infertility. wait for results. (ii)PGT-A can still miss certain issues like mosaicism where an embryo has a mix of normal and abnormal cells, and errors can

					occur due to limitations in trophectoderm biopsy procedures
van de Wiel et al.,(2020)	Observational	TLI and PGT	NA	NA	TLI is a relatively inexpensive add-on, with prices ranging from £0 to £795 and PGT A usually costs between £2,100 and £3,295, with an average price of £2,695.
Masbou et al., (2019)	Retrospective	PGT-A	-	(i)Implantation Rate: 63.2% (ii)Live Birth Rate: 54% (iii)Miscarriage Rate: 14.5%	PGT is an invasive procedure can lead to cellular damage, potentially affecting the embryo's development and its implantation potential
Ozgur et al., (2019)	RCT	PGT-A with NGS	(i)Fertilization Rate: 81.0% (ii)Blastocyst formation Rate: 53.0%	Patient under 35: (i)Pregnancy Rate: 61.3% (ii)Miscarriage Rate: 6.1% (iii)Live Birth Rate: 58.6%	Time for PGT results takes 7 days after the biopsy.
Barberet et al., (2018)	RCT	TLI	(i)Fertilization rate: 75.4% (ii)Blastocyst formation rate: 44.8% (iii)Viable embryo rate: 47.8%	(i)Implantation Rate: 35.1% (ii)Pregnancy Rate: 48.6% (iii)Miscarriage rate: 14.9%	By day 3 or day 5, embryologists have a full developmental profile to inform selection and transfer
Verpoest et al., (2018)	RCT	PGT-A with aCGH	–	(i)Pregnancy Rate: 38% (ii)Miscarriage Rate: 8% (iii)Live Birth Rate: 30%	PGT involves advanced procedures such as laser polar body biopsy and array comparative genomic hybridization requires specialized lab equipment and reagent such as those for DNA extraction, amplification, sequencing and trained personnel.
Goodman et al., (2016)	RCT	TLI with KID score	(i)Fertilization rate: 57.24% (ii)KIDscore TLI (2.6 ± 1.2)	(i)Implantation Rate: 51% (ii)Pregnancy Rate: 68%	TLI allows the identification of multi nucleation and irregular cell divisions, which are linked to reduced implantation success.
Kieslinger et al., (2016)	Prospective	TLI with EEVA	(i)Fertilization rate: 67.4% (ii)Top quality embryo: 82.3%	(i)Implantation Rate: 35% (ii)Pregnancy Rate: 40.4% (iii)Live Birth Rate: 32.9%	Embryos remain in a stable incubator environment, avoiding potential stress from handling and environmental changes.
Siristatidis et al., (2015)	Prospective	TLI	(i)Fertilization Rate: 65.4%	(i)Pregnancy Rate: 65.7% (ii)Live Birth Rate: 45.71%	TLI helps predict which embryos will reach the blastocyst stage, optimizing transfer timing and improving outcomes for patients.

Park et al., (2015)	RCT	TLI	(i)Fertilization rate: 49.47% Mean good quality embryo (2.41±2.3)	(i)Implantation Rate: 27.9% (ii)Pregnancy Rate: 33.5% (iii)Miscarriage Rate: 33.3%	-
Fodina et al., (2015)	Cohort	TLI	Fertilization Rate: 64.4%	Total all clinical outcome rates: (i)Implantation rate: 31.7% (ii)Pregnancy Rate: 39.7% (iii)Miscarriage Rate: 12.1% Patient younger than 35: (i)Implantation rate: 35.5% (ii)Pregnancy Rate: 39.8% (iii)Miscarriage Rate: 0% Patient older than 35: (i)Implantation rate: 15.8% (ii)Pregnancy Rate: 27.8% (iii)Miscarriage Rate: 10.0%	Non-invasive and stable culture conditions, minimizes exposure to fluctuations in CO_2, pH, temperature due to no need to remove embryos from incubator.
Rubio et al., (2014)	Prospective	TLI	(i)Fertilization Rate: 75.3% (ii)Blastocyst formation rate: 52.3%	(i)Pregnancy Rate: 65.3% (ii)Miscarriage rate: 16.6%	The system of TLI has precise timing of embryonic development helps reduce observer subjectivity

*Note. TLI = Time-Lapse Imaging; PGT = Preimplantation Genetic Testing; NA: Not applicable: PGT-A = Preimplantation Genetic Testing for Aneuploidy; RCT = Randomized Controlled Trial; EEVA = Early Embryo Viability Assessment; NGS = Next-Generation Sequencing; SNP = Single- Nucleotide Polymorphism; aCGH = Array Comparative Genomic Hybridization; iDA = Intelligent Data Analysis; KID = Kinetics Information Data

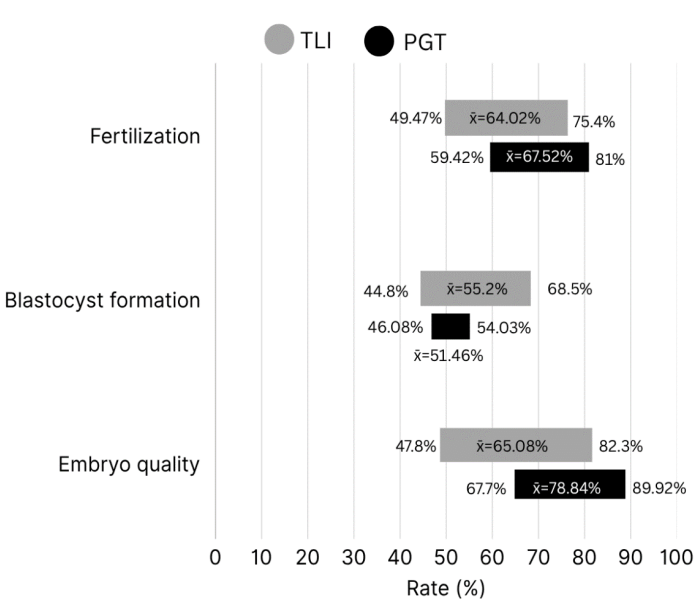


Figure 2: Horizontal Bar Chart comparing TLI and PGT effectiveness in selecting the most viable embryos in terms of fertilization rate, blastocyst formation rate and embryo quality

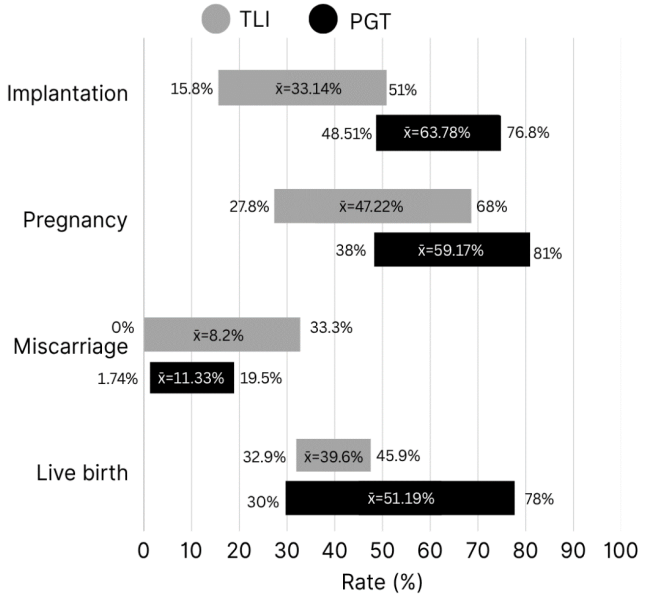


Figure 3: Horizontal Bar Chart comparing TLI and PGT effectiveness in clinical outcomes in terms implantation rate, pregnancy rate, miscarriage rate and live birth rate

Table 3: Strength and limitation of TLI and PGT in terms of time, invasiveness, accuracy and cost

Aspect	TLI	PGT
Time	• Very quick takes only 21 seconds to check embryos (Illingworth et al., 2024). • By day 3 to 5, gives a full growth profile for transfer (Barberet et al., 2018). • Can predict which embryos reach blastocyst stage (Siristatidis et al., 2015).	• Slower because includes biopsy, genetic testing, freezing, and counselling (Sui et al., 2020). • Results usually take around 7 days (Ozgur et al., 2019).
Invasiveness	• Non-invasive as embryos stay in incubator, no disturbance (Fodina et al., 2015). • Stable culture, no stress from handling (Kieslinger et al., 2016).	• Invasive because embryo biopsy may cause cell damage (Masbou et al., 2019). • Many cycles fail to reach transfer due to embryo loss (Sui et al., 2020).
Accuracy	• Detects abnormal cell divisions linked to poor outcomes (Goodman et al., 2016). • Tracks exact timing of divisions, reducing subjectivity (Rubio et al., 2014). • Spots problems like direct/reverse cleavage in real-time (Meng et al., 2022).	• Next generation sequencing gives detailed chromosome analysis (Singh et al., 2022). • May miss mosaicism (Sui et al., 2020). • Errors possible from biopsy limitations.
Cost	• Cheaper around £0 to £795 with median £478(van de Wiel et al., 2020). • Requires monitoring and recording of the embryos over time, which is less labor-intensive and require fewer human resources (Bhide et al., 2024).	• Much more expensive around £2,100 to £3,295 with median £2,695 (van de Wiel et al., 2020). • Full package may cost more £9,500 (van de Wiel et al., 2020). • Requires advanced labs & procedures (Verpoest et al., 2018).

DISCUSSSION

Effectiveness of TLI and PGT in selecting the most viable embryos for transfer, in terms of fertilization rate, blastocyst formation rate, and embryo quality

PGT generally demonstrates a better performance in fertilization rate and embryo quality. The mean fertilization rate for PGT is 67.52% (Harris et al., 2024; Katz-Jaffe et al., 2024; Zheng et al., 2023; Singh et al., 2022; Sui et al., 2020; Ozgur et al., 2019), which is higher than TLI's 64.02% (Bhide et al., 2024; Illingworth et al.,2024; Barberet et al., 2018; Goodman et al., 2016; Kieslinger et al., 2016; Fodina et al., 2015; Park et al., 2015; Siristatidis et al., 2015; Rubio et al., 2014), and it quite falls within the normal range of fertlization rate for this procedure (Thamarai Fertility, 2018). While some studies show modest improvements with PGT (Sui et al., 2020), others report more significant gains (Singh et al., 2022), indicating that PGT offers more stable and predictable fertilization outcomes compared to the widely varying results seen with TLI. Additionally, PGT shows a higher average embryo quality score of 78.84% (Xu et al., 2024; Zheng et al., 2023) compared to TLI's 65.08% (Barberet et al., 2018; Kieslinger et al., 2016). This can be attributed to several factors, including the fact that patients undergoing PGT may be in more advanced IVF centers with highly skilled embryologists and optimized lab conditions.

When it comes to blastocyst formation rate, TLI shows a slight advantage with an average of 55.2% (Illingworth et al.,2024; Barberet et al., 2018; Rubio et al., 2014) compared to PGT's 51.46% (Harris et al., 2024; Singh et al., 2022; Ozgur et al., 2019). The blastocyst formation rate is the percentage of fertilized eggs that develop into blastocysts (Day 5 or 6 embryos), and the the success of PGT in this area is highly dependent on maternal age, with some younger women under 35 years old showing a higher blastocyst formation rate of 54.03% (Harris et al., 2024). It's also important to note that a cohort of some embryos intended for PGT will inherently contain aneuploid or mosaic embryos, which can influence the overall observed blastocyst formation rate. In contrast, TLI has shown the potential for more significant improvements in blastocyst development in specific contexts (Illingworth et al., 2024).

Impact of TLI and PGT on clinical outcomes, including implantation rates, pregnancy rates, miscarriage rates, and live birth rate

PGT consistently demonstrates superior clinical outcomes compared to TLI across several critical aspects. For instance, PGT's average implantation rate is 63.78% (Katz-Jaffe et al., 2024; Singh et al., 2022; Sui et al., 2020; Masbou et al., 2019; Ozgur et al., 2019), significantly higher than TLI's 33.14% (Barberet et al., 2018; Goodman et al., 2016; Kieslinger et al., 2016; Fodina et al., 2015; Park et al., 2015). This pronounced difference is largely attributed to PGT's ability to identify and select genetically healthy, or euploid, embryos, which directly addresses a primary cause of implantation failure. Similarly, PGT shows a higher average pregnancy rate of 59.17% (Harris et al., 2024; Xu et al., 2024; Zheng et al., 2023; Singh et al., 2022; Sui et al., 2020; Verpoest et al., 2018) compared to TLI's 47.22% (Bhide et al., 2024; Illingworth et al.,2024; Meng et al., 2022; Barberet et al., 2018; Goodman et al., 2016; Kieslinger et al., 2016; Park et al., 2015; Fodina et al., 2015; Siristatidis et al., 2015; Rubio et al., 2014) and a greater likelihood of a live birth, with a mean rate of 51.19% in PGT (Harris et al., 2024; Katz-Jaffe et al., 2024; Xu et al., 2024; Zheng et al., 2023; Singh et al., 2022; Sui et al., 2020; Masbou et al., 2019; Ozgur et al., 2019; Verpoest et al., 2018) versus TLI's 39.60% (Bhide et al., 2024; Illingworth et al., 2024; Meng et al., 2022; Kieslinger et al., 2016; Siristatidis et al., 2015). While TLI selects morphologically normal embryos, it does not evaluate genetic health, a key factor in achieving ongoing pregnancy, which may explain its lower success rates.

Conversely, while PGT excels in the majority of success aspects, TLI shows a slightly lower average miscarriage rate. TLI studies reported an average miscarriage rate of 8.2% (Bhide et al., 2024; Meng et al., 2022; Barberet et al., 2018; Fodina et al., 2015; Park et al., 2015; Rubio et al., 2014), whereas PGT studies had an average of 11.33% (Harris et al., 2024; Katz-Jaffe et al., 2024; Xu et al., 2024; Singh et al., 2022; Sui et al., 2020; Masbou et al., 2019; Ozgur et al., 2019; Verpoest et al., 2018). This unexpected finding may be due to the differing patient populations studied; PGT studies is often utilized for patients with a higher baseline risk of miscarriage, such as those with advanced maternal age or a history of recurrent pregnancy loss, which could skew the average higher despite PGT's effectiveness in preventing genetic miscarriages. PGT's targeted approach of screening for chromosomal abnormalities is highly effective, leading to a reduction in miscarriage rates in many studies (Singh et al., 2022; Xu et al., 2024), but this benefit is not universally reported, which contributes to the observed range and average.

Strengths and limitations of TLI and PGT in terms of cost, invasiveness, time, and accuracy

TLI has high efficiency, non-invasive nature, and cost-effectiveness. The process is rapid, with embryo assessment taking as little as 21.3 seconds using Intelligent Data Analysis (iDA) algorithms (Illingworth et al., 2024). This

non-invasive approach avoids direct embryo manipulation, minimizing the risk of damage and maintaining a stable culture environment, which is crucial for embryo integrity (Kieslinger et al., 2016). Furthermore, TLI is more affordable, with costs ranging from £0 to £795, making it a more accessible option for many patients (van de Wield et al., 2020). However, TLI's primary limitation is its inability to directly assess genetic health. While it can identify morphological abnormalities like multi-nucleation and irregular cell divisions, embryos with ideal appearance may still carry genetic defects such as aneuploidy, which can compromise their viability despite appearing healthy (Meng et al., 2022).

In contrast, PGT excels in accuracy and genetic analysis but is limited by its invasiveness, cost, and time-consuming nature. PGT utilizes advanced techniques such as next-generation sequencing and array comparative genomic hybridization to provide a comprehensive genetic profile of the embryo, enabling the detection of chromosomal abnormalities like Down syndrome (Singh et al., 2022). This high level of genetic screening is crucial for patients with known hereditary risks or advanced maternal age, as it significantly reduces the risk of miscarriage and improves the likelihood of a successful pregnancy. The procedure is invasive, requiring a blastocyst biopsy, which can potentially damage the embryo and result in a failure to transfer in some cases (Sui et al., 2020). The process is also time-intensive, with results taking about 7 days to become available, and it is significantly more expensive, with costs ranging from £2,100 to £9,500 depending on the package (van de Wield et al., 2020). While PGT offers superior genetic insight, it may also yield false positives or negatives due to mosaicism, where some cells are normal and others are not, potentially leading to the discarding of viable embryos (Sui et al., 2020)

Limitation of study

A major weakness is the absence of direct, head-to-head comparative studies between TLI and PGT, forcing the review to evaluate each technology independently. This approach may not accurately reflect their combined effectiveness in clinical practice. The study's conclusions are also limited by the significant variability across the included literature, including differences in protocols, patient populations, and success metrics. The inconsistent nature of embryo assessment methods further complicates direct comparisons. Additionally, the review is susceptible to selection bias, as differences in patient populations also contribute to this bias which TLI is often used in broader, healthier populations, while PGT is typically reserved for cases with a poorer prognosis, like advanced maternal age. Finally, the review notes that PGT studies may report more favorable outcomes because they are often conducted in highly specialized IVF centers with more skilled staff and optimized lab conditions, which is a factor not consistently present in TLI studies.

Future research directions

Future research on embryo selection for in vitro fertilization should focus on direct comparisons between TLI and PGT to better understand their combined effectiveness. While existing studies often examine these technologies in isolation, or focus on how they complement each other, there's a critical need for large-scale, methodologically rigorous studies, such as randomized controlled trials, to provide more definitive evidence on their performance in clinical settings. Furthermore, the integration of artificial intelligence is rapidly advancing both TLI and PGT, with AI-driven algorithms analyzing time-lapse data to predict embryo viability and ploidy (Barnes et al., 2023). Similarly, non-invasive PGT is revolutionizing genetic screening by analyzing cell-free DNA, potentially eliminating the need for a biopsy and reducing associated risks (Cengiz Cinnioğlu et al., 2023). These technological advancements make direct comparative studies even more crucial for refining assisted reproductive practices and improving patient outcomes.

CONCLUSION

This scoping review highlights TLI and PGT are two distinct but complementary tools used in ART to improve embryo selection. TLI provides a non-invasive, continuous assessment of embryo development, which is particularly effective for predicting blastocyst formation and improving embryo quality. However, its impact on overall pregnancy and live birth rates is modest. In contrast, PGT is a more invasive but highly effective genetic screening method that significantly improves live birth rates by identifying and preventing the transfer of chromosomally abnormal embryos, which is especially beneficial for older patients or those with a poor prognosis. The decision to use either technology depends on individual patient factors, such as the need for genetic screening and the balance of associated costs and risks.

ACKNOWLEDGEMENT

We would like to express our gratitude to the Department of Biomedical Science, Kulliyyah of Allied Health Sciences, IIUM, for their invaluable support and motivation. This research was not funded by any grant.

REFERENCES

Barberet, J., Chammas, J., Bruno, C., Valot, E., Vuillemin, C., Jonval, L., Choux, C., Sagot, P., Soudry, A., & Fauque, P. (2018). Randomized controlled trial comparing embryo culture in two incubator systems: G185 K-System versus EmbryoScope. Fertility and Sterility, 109(2), 302-309.e1.https://doi.org/10.1016/j.fertnstert.2017.10.008

Barnes, J., Brendel, M., Gao, V. R., Rajendran, S., Kim, J., Li, Q., Malmsten, J. E., Sierra, J. T., Zisimopoulos, P., Sigaras, A., Khosravi, P., Meseguer, M., Zhan, Q., Rosenwaks, Z., Elemento, O., Zaninovic, N., & Hajirasouliha, I. (2023). A non-invasive artificial intelligence approach for the prediction of human blastocyst ploidy: a retrospective model development and validation study. The Lancet Digital Health, 5(1), e28–e40. https://doi.org/10.1016/S2589-7500(22)00213-8

Bhide, P., Chan, D. Y. L., Lanz, D., Odai Alqawasmeh, Barry, E., Baxter, D., Carreras, F. G., Choudhury, Y., Cheong, Y., Pui, J., Collins, B., Cong, L., Doidge, S., 64 Heighway, J., Patel, D., Pardo, M. C., Rattos, A., Wright, A., Dodds, J., & Perez, T. (2024). Clinical effectiveness and safety of time-lapse imaging systems for embryo incubation and selection in in-vitro fertilisation treatment (TILT): a multicentre, three-parallel-group, double-blind, randomised controlled trial. The Lancet, 404(10449), 256–265. https://doi.org/10.1016/s0140-6736(24)00816-x

Burks, C., Van Heertum, K., & Weinerman, R. (2021). The technological advances in embryo selection and genetic testing: a look back at the evolution of aneuploidy screening and the prospects of non-invasive PGT. Reproductive Medicine, 2(1), 26–34. https://doi.org/10.3390/reprodmed2010004

Cengiz Cinnioğlu, Glessner, H., Jordan, A. S., & Bunshaft, S. (2023). A systematic review of noninvasive preimplantation genetic testing for aneuploidy. Fertility and Sterility, 120(2), 235–239. https://doi.org/10.1016/j.fertnstert.2023.06.013

Fodina, V., Andzane, D., Evija Pimane, & Miskova, A. (2015). How to decrease EPL rate treating infertility? Gynecological Endocrinology, 31(sup1), 51–54. https://doi.org/10.3109/09513590.2015.1086508

Goodman, L. R., Goldberg, J., Falcone, T., Austin, C., & Desai, N. (2016). Does the addition of time-lapse morphokinetics in the selection of embryos for transfer improve pregnancy rates? A randomized controlled trial. Fertility and Sterility, 105(2), 275-285.e10. https://doi.org/10.1016/j.fertnstert.2015.10.013

Harris, B. S., Acharya, K. S., Shakthi Unnithan, Neal, S. A., Mebane, S., Truong, T., & Muasher, S. J. (2024). Success rates with preimplantation genetic testing for aneuploidy (PGT-A) in good prognosis patients is dependent on age. Fertility and Sterility. https://doi.org/10.1016/j.fertnstert.2024.09.043

Illingworth, P. J., Venetis, C., Gardner, D. K., Nelson, S. M., Berntsen, J., Larman, M. G., Agresta, F., Ahitan, S., Ahlström, A., Cattrall, F., Cooke, S., Demmers, K., Gabrielsen, A., Hindkjær, J., Kelley, R. L., Knight, C., Lee, L., Lahoud, R., Mangat, M., & Park, H. (2024). Deep learning versus manual morphology-based embryo selection in IVF: a randomized, double-blind noninferiority trial. Nature Medicine. https://doi.org/10.1038/s41591-024-03166-5

Katz-Jaffe, M., Gassen, C., Makloski, R., Reed, L., & Schoolcraft, W. B. (2024). Impact of aneuploidy on reproductive success in young infertile women: prospective analysis. Reproductive Biomedicine Online, 103858–103858. https://doi.org/10.1016/j.rbmo.2024.103858

Kieslinger, D. C., De Gheselle, S., Lambalk, C. B., De Sutter, P., Kostelijk, E. H., Twisk, J. W. R., van Rijswijk, J., Van den Abbeel, E., & Vergouw, C. G. (2016). Embryo selection using time-lapse analysis (Early Embryo Viability Assessment) in conjunction with standard morphology: a prospective two-center pilot study. Human Reproduction, 31(11), 2450–2457. https://doi.org/10.1093/humrep/dew207

Kupka, M. S., Chambers, G. M., Dyer, S., Zegers-Hochschild, F., Jacques de Mouzon, Ishihara, O., Banker, M., Seung Chik Jwa, Fu, B., Eman Elgindy, Baker, V., & G David Adamson. (2024). International Committee for Monitoring Assisted Reproductive Technology world report: assisted reproductive technology, 2015 and 2016. Fertility and Sterility. https://doi.org/10.1016/j.fertnstert.2024.07.009

Masbou, A. K., Friedenthal, J. B., McCulloh, D. H., McCaffrey, C., Fino, M. E., Grifo, J. A., & Licciardi, F. (2019). A comparison of pregnancy outcomes in patients undergoing donor egg single embryo transfers with and without preimplantation genetic testing. Reproductive Sciences, 26(12), 1661–1665. https://doi.org/10.1177/1933719118820474

Meng, Q., Xu, Y., Zheng, A., Li, H., Ding, J., Xu, Y., Pu, Y.,

Wang, W., & Wu, H. (2022). Noninvasive embryo evaluation and selection by time-lapse monitoring vs. conventional morphologic assessment in women undergoing in vitro fertilization/intracytoplasmic sperm injection: a single-center randomized controlled study. Fertility and Sterility, 117(6), 1203–1212. https://doi.org/10.1016/j.fertnstert.2022.02.015

Ozgur, K., Berkkanoglu, M., Bulut, H., Yoruk, G. D. A., Candurmaz, N. N., & Coetzee, K. (2019). Single best euploid versus single best unknown-ploidy blastocyst frozen embryo transfers: a randomized controlled trial. Journal of Assisted Reproduction and Genetics, 36(4), 629–636. https://doi.org/10.1007/s10815-018-01399-1

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., racowJ., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical Research ed.), 372, n71. https://doi.org/10.1136/bmj.n71

Papamentzelopoulou, M., Prifti, I., Mavrogianni, D., Tseva, T., Soyhan, N., Athanasiou, A., Athanasiou, A., Athanasiou, A., Vogiatzi, P., Konomos, G., Loutradis, D., & Sakellariou, M. (2024). Assessment of artificial intelligence model and manual morphokinetic annotation system as embryo grading methods for successful live birth prediction: a retrospective monocentric study. Reproductive Biology and Endocrinology, 22(1). https://doi.org/10.1186/s12958-024-01198-7

Park, H., Bergh, C., Selleskog, U., Thurin-Kjellberg, A., & Lundin, K. (2015). No benefit of culturing embryos in a closed system compared with a conventional incubator in terms of number of good quality embryos: results from an RCT. Human Reproduction, 30(2), 268–275. https://doi.org/10.1093/humrep/deu316

Rubio, I., Galán, A., Larreategui, Z., Ayerdi, F., Bellver, J., Herrero, J., & Meseguer, M. (2014). Clinical validation of embryo culture and selection by morphokineticanalysis: a randomized, controlled trial of the EmbryoScope. Fertility and sterility, 102(5), 1287-1294.e5. https://doi.org/10.1016/j.fertnstert.2014.07.738

Singh, P., Pathak, A., Singh, P., & Desai, M. (2022). Selecting euploid embryos for transfer by preimplantation genetic testing with the help of next-generation sequencing in poor prognosis patients: A retrospective cohort analysis. 15(2), 157–157. https://doi.org/10.4103/jhrs.jhrs_166_21

Siristatidis, C., Komitopoulou, M. A., Makris, A., Sialakouma, A., Botzaki, M., Mastorakos, G., Salamalekis, G., Bettocchi, S., & Palmer, G. A. (2015). Morphokinetic parameters of early embryo development via time lapse monitoring and their effect on embryo selection and ICSI outcomes: a prospective cohort study. Journal of Assisted Reproduction and Genetics, 32(4), 563–570. https://doi.org/10.1007/s10815-015-0436-z

Sravya Gudapati, Chaudhari, K., Shrivastava, D., & Yelne, S. (2024). Advancements and applications of preimplantation genetic testing in in vitro fertilization: a comprehensive review. Curēus. https://doi.org/10.7759/cureus.57357

Sui, Y., Lei, C., Ye, J., Fu, J., Zhang, S., Li, L., Peng, X., Zhang, Y., Chen, G., & Sun, X. (2020). In vitro fertilization with single-nucleotide polymorphism microarray-based preimplantation genetic testing for aneuploidy significantly improves clinical outcomes in infertile women with recurrent pregnancy loss: A randomized controlled trial. Reproductive and Developmental Medicine, 4(1), 32–41. https://doi.org/10.4103/2096-2924.281852

Thamarai Fertility. Fertilization and Embryo Culture (2018). Thamaraihealthcare.com. https://thamaraihealthcare.com/fertilization-and-embryo-culture/

van de Wiel, L., Wilkinson, J., Athanasiou, P., & Harper, J. (2020). The prevalence, promotion and pricing of three IVF add-ons on fertility clinic websites. Reproductive BioMedicine Online, 41(5), 801–806. https://doi.org/10.1016/j.rbmo.2020.07.021

Verpoest, W., Staessen, C., Bossuyt, P. M., Goossens, V., Altarescu, G., Bonduelle, M., Devesa, M., Eldar-Geva, T., Gianaroli, L., Griesinger, G., Kakourou, G., Kokkali, G., Liebenthron, J., Magli, M.-C., Parriego, M., Schmutzler, A. G., Tobler, M., van der Ven, K., Geraedts, J., & Sermon, K. (2018). Preimplantation genetic testing for aneuploidy by microarray analysis of polar bodies in advanced maternal age: a randomized clinical trial. Human Reproduction, 33(9), 1767–1776. https://doi.org/10.1093/humrep/dey262

Wang, J., Guo, Y., Zhang, N., & Li, T. (2023). Research progress of time-lapse imaging technology and embryonic development potential: A review. Medicine, 102(38), e35203–e35203. https://doi.org/10.1097/md.0000000000035203

Xu, N., Du, Y., Su, Z., Yu, S., Li, W., Guan, Y., & Lou, H. (2024). Analysis of pregnancy outcomes in patients with unexplained recurrent miscarriage assisted by IVF/ICSI

with or without PGT-A. Journal of Assisted Reproduction and Genetics. https://doi.org/10.1007/s10815-025-03438-0

Zheng, Z., Tan, J., Chen, L., Liu, S., Zhou, C., & Li, Y. (2023). PGT-A improved singleton live birth rate among all age groups of women who underwent elective single blastocyst transfer: a single-centre retrospective study. Journal of Assisted 73 Reproduction and Genetics, 40(6), 1417–1427. https://doi.org/10.1007/s10815-023-02775-2

REVIEW ARTICLE

Anticancer Properties of Bee Venom Bioactive Compounds: A Scoping Review

Siti Nur Farisyah Abdul Rahim[1], Mohammad Syaiful Bahari Abdull Rasad[1,*]

[1]Department of Biomedical Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
Bee venom; anticancer activity; melittin; bioactive compounds

ABSTRACT

Background: *Apis melliferra* (bee venom) has long been used in alternative medicine for its therapeutic effects and recently research highlights its anticancer potential. This scoping review aims to systematically gather and evaluate evidence on the anticancer properties of bee venom and its bioactive compounds. **Methods:** This review followed PRISMA-ScR guidelines using articles retrieved from PubMed, ScienceDirect and Scopus from 2010 to 2024. Studies were screened based on inclusion and exclusion criteria, focusing on the anticancer effects of bee venom components. A total of 240 articles were retrieved, with 35 meeting the eligibity criteria. Data on cancer type, model, compound, and mechanism were extracted and analyzed. **Results:** It was found that melittin was the most frequently investigated compound, showing anticancer activity in breast, lung, glioblastoma, and ovarian cancer models. Mechanisms include induction of apoptosis, inhibition of cell proliferation, suppression of NFκB and PI3K/Akt/mTOR pathways, and modulation of immune responses. Additionally, a number of studies documented synergistic effects when melittin was mixed with apamin and PLA2, two additional components of bee venom. **Conclusion:** Bee venom, particularly melittin, demonstrated strong anticancer potential by targeting multiple molecular pathways. However, further *in vivo* and clinical studies are required to validate its efficacy and safety as a therapeutic agent in addition to thorough examination of less well-known substances such as adolapin and MCD peptide.

INTRODUCTION

Nearly 20 million new cases of cancer in 2022 with 9.7 million deaths from cancer and its global incidence continues to rise despite major advances in diagnostics and therapeutics (Bray et al., 2024). Many conventional cancer treatments including surgery, chemotherapy and radiotherapy, often present significant limitations such as severe side effects, multidrug resistance and non-selective toxicity toward healthy issues. These challenges highlight the pressing need for new anticancer agents that are both selective and less harmful. In recent decades natural bioactive compounds derived from plants, marine organisms and animal venoms have gained attention as potential sources of novel anticancer drugs due to their ability to target multiple signaling pathways involved in tumour development and metastasis (Zhang et al., 2018). Among naturally derived agents, bee venom extracted from a bee species known as *Apis mellifera* has attracted considerable interest for its broad spectrum of pharmacological properties, including anti-inflammatory, antimicrobial, and also anticancer effects. Traditionally, bee venom therapy has been practiced for centuries in various cultures such as Chinese, Indian, and Malay medicine for the treatment of arthritis, pain, and inflammatory disorders (Pandey et al., 2023). Modern applications of bee venom, often referred to as apitherapy, have expanded into the management of chronic diseases, autoimmune conditions, and cancer. Bee venom is a complex mixture of peptides, enzymes, and biogenic amines, with melittin, phospholipase A2 (PLA2), apamin, adolapin, and mast cell degranulating (MCD) peptide as its principal bioactive components (Abd El-Wahed et al., 2019). Among these, melittin comprising approximately 40-60% of the venom's dry weight is the most studied and exhibits potent anticancer activity across a wide range of cancer types, including breast, lung, ovarian, liver, and glioblastoma. Melittin exerts its effects by inducing apoptosis, suppressing tumour cell proliferation, disrupting cellular membranes, and modulating key oncogenic signaling pathways such as PI3K/Akt and NF-κB (Wehbe et al., 2019). PLA2 facilitates melittin's entry into tumour cells and enhances its cytolytic effects, while adolapin and apamin exhibit anti-inflammatory and neuroprotective properties that may complement bee venom's anticancer potential (Cui et al., 2024).

At therapeutic dosages, several investigations have shown that bee venom and its constituents can specifically target cancerous cells while preserving healthy ones. However,

* *Corresponding author.*
E-mail address: syaiful@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

concerns about toxicity, delivery effectiveness, and a lack of knowledge about its molecular mechanisms continue to restrict the medicinal potential of bee venom, even in the face of encouraging preclinical results. Additionally, most of the research has concentrated on melittin, leaving other substances like adolapin and MCD peptide comparatively unexplored (Pandey et al., 2023).

The current scoping review fills these gaps by methodically mapping and synthesising scientific data in the anticancer effects of bee venom and its bioactive ingredients that were published between 2010 and 2024. This review aims to identify the mechanisms of action, summarize experimental findings across cancer models, and highlight research gaps to guide future exploration of bee venom-derived compounds as potential anticancer therapeutics. This review may represent the most up-to-date and comprehensive analysis of bee venom components and their anticancer properties.

METHODOLOGY

Study Design

This study adopted a scoping review design based on the methodological framework proposed by Arksey & O'Malley (2005) guided by the Prisma-ScR checklist, comprising five key stages: (1) identifying research questions, (2) identifying relevant studies, (3) study selection, (4) charting the data, and (5) collating, summarizing, and reporting the results (Arksey & O'Malley, 2005). The scoping approach was chosen to map and summarize the existing literature regarding the anticancer potential of bee venom and its bioactive components, including melittin, phospholipase A2 (PLA2), apamin, mast-cell degranulating (MCD) peptide, and hyaluronidase.

Search Strategy

A comprehensive and systematic search strategy was employed to retrieve relevant studies. Three electronic databases, PubMed, ScienceDirect, and Scopus were searched for publications between 2010 and 2024. The keywords were constructed using Medical Subject Headings (MeSH) terms and Boolean operators (AND, OR) are used to maximize sensitivity and specificity. The search terms included combinations of "bee venom", "anticancer activity", "bioactive compound" and "antineoplastic agents" were included. The search was limited to English language, peer-reviewed articles. All retrieved were exported into Microsoft Excel for organization and duplicate removal.

Inclusion and Exclusion Criteria

The inclusion criteria: (1) Studies investigating the anti-cancer effects and mechanisms of actions of bee venom or its bioactive compound, (2) full text articles, (3) peer-reviewed articles published in English from 2010 until 2024, and (4) experimental studies conducted *in vitro, in vivo,* or clinical studies involving both human and animal studies. Studies were excluded if the studies focus on non-bee venom toxins or those unrelated to cancer, articles without full text access, published before 2010 and after 2024, systematic reviews, meta-analyses, and other reviews lacking primary data or original research. These criteria ensured that only relevant and high-quality studies were included to provide reliable evidence on the anticancer mechanisms of bee venom.

Study Selection

The selection process followed the PRISMA-ScR guidelines. An initial total of 240 records were identified from the three selected databases. After removing 72 duplicates, 168 studies remained for title and abstract screening. Following this stage, 65 full-text articles were assessed for eligibility, and 35 met all inclusion criteria and were included in the final review. Screening was performed in two stages: (1) title and abstract screening to exclude irrelevant studies, (2) full-text screening to confirm eligibility based on inclusion and exclusion criteria. Discrepancies during screening were resolved through discussion to ensure consistency. The detailed process is illustrated in **Figure 1**, following the PRISMA-ScR flow structure.

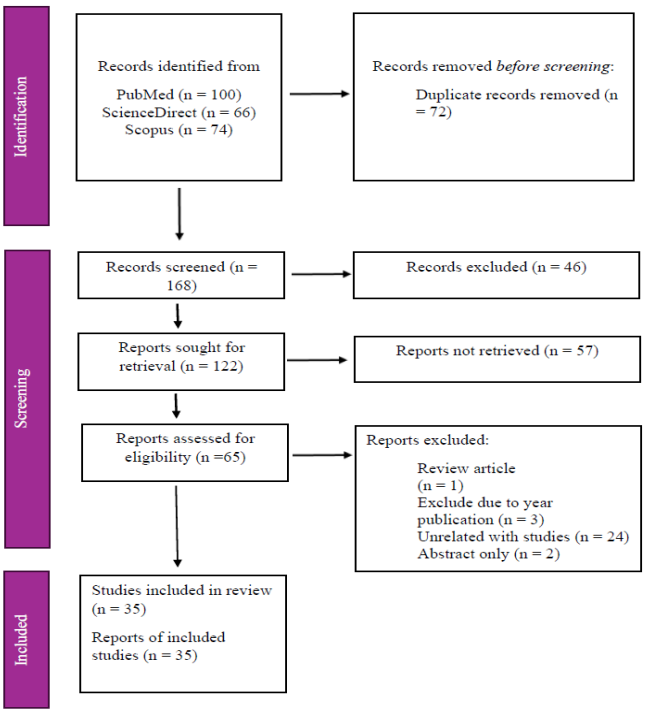


Figure 1: PRISMA-ScR flow diagram

Data Extraction and Charting

Data extraction and charting were performed systematically and summarized in the table form from the final full-text articles and journals. For each included article, the following information was extracted and organized: (1) author(s) and year of publication, (2) cancer type, (3) cancer model, (4) bioactive compound, (5) IC_{50}/LC_{50} values, and (6) reported anticancer activity from the study. Table 1 summarizes the data, presenting an overview of the included studies and highlighting the primary mechanisms through which bee venom and its components exert anticancer effects.

Table 1: Extraction of data of the bioactive compounds of bee venom and their anticancer effects.

Cancer Type	Cancer Model	Bioactive Compound	IC_{50}/LC_{50}	Reported Anticancer Activity	Ref.
Glioblastoma	8-MG-BA, GAMG (human glioblastoma cell lines)	• Melittin • PLA2 • Apamin	Not specified	Inhibit cell viability, inhibit MMP-2 and MMP-9 secretion (anti-metastatic)	Malek et al., 2022
	U87 human glioblastoma cells in nude mice	• Apamin • Melittin • PLA2 • MCD-peptide	IC_{50}: 14.32 µg/mL (melittin *in vitro*)	Apoptosis (EGFP+/Caspase-3+), necrosis	Chahla et al., 2024
	A127 (human glioblastoma cell lines)	• Melittin • PLA2	IC_{50}: 28.5 µg/mL	Induced apoptosis, inhibits MMP-2 expression and activity	Sisakh et al., 2017
Chronic Myelogenous Leukaemia (CML)	K562 cells (human leukaemia cell line)	• Melittin	IC_{50}: 1.84 µg/mL	Induced apoptosis, induced necrosis, and cell cycle arrest at G0/G1 phase	Obeidat et al., 2023
Acute Lymphoblastic Leukaemia (ALL)	CCRF-CEM (Human leukaemia cell lines)	• Melittin	Not specified	Induced apoptosis	Ceremuga et al., 2020
Breast cancer	SUM159	• Melittin	IC_{50}: 0.94 µM	Induced apoptosis, inhibits EGFR/HER2 activation	Duffy et al., 2020
	MDA-MB-231	• Melittin	IC_{50}: 5.86 µg/mL	Inhibit metastasis, suppression of SDF-1α/CXCR4 signalling cell, induces apoptosis	Salimian et al., 2022
	MCF-7 (*in vitro*)	• Melittin • PLA2 • Apamin	IC_{50}: 1.8 µg/mL	Induced apoptosis, increase p53 and 8-OHdG levels, inhibit tumour growth	Plasay et al., 2022
				Inhibits metastasis, downregulating CD147 and MMP-9 expression	Shaw et al., 2019, Wang et al., 2017
				Inhibit NF-kB pathway	Kim et al., 2015
Cervical cancer	HeLa	• Melittin	IC_{50}: 1.8 µg/mL	Induced apoptosis, increase p53 and 8-OHdG	Zarrinnahad et al., 2017

				levels	
	C33A	• Melittin	IC_{50}: 21.59-23.12 µg/mL	Cytotoxicity, ROS generation	Shasuvar et al., 2023
Malignant melanoma	B16F10	• Melittin	IC_{50}: 12.67 µg/mL	Inhibit cell growth, inhibits migration, induced apoptosis	Lim et al., 2019
Colon cancer	HT-29 (*in vitro*)	• Melittin • PLA2 • Apamin	IC_{50}: 5.86 µg/mL	Induced apoptosis, inhibit proliferation	Duarte et al., 2022
	HCT116	• Melittin • PLA2	IC_{50}: 101.0-194.9 µg/mL	Cytotoxicity, synergistic activity of melittin and PLA2, inhibits proliferation	Yaacoub et al., 2021
		• Melittin	IC_{50}: 3.8 µg/mL	Induced apoptosis, activates autophagy pathways	El-Didamony et al., 2024
			IC_{50}: 14.05 µg/mL	Cytotoxicity	Zamani et al., 2024
	COLO205	• Melittin	Not specified	Induced apoptosis	Soliman et al., 2019
	HCT-15	• Melittin	Not specified	Induced apoptosis	Soliman et al., 2019
Gastric cancer	AGS (Human gastric cancer cell line)	• Melittin	Not specified	Inhibits cell viability, Induced necrosis at higher dose	Mahmoodzadeh et al., 2015
		•		Induced apoptosis, inhibit proliferation	Soliman et al., 2019
Hepatocellular carcinoma cancer (HCC)	HepG2	• Melittin	Not specified	Inhibit proliferation, inhibit migration, induces demethylation of ADAMTS9-AS2	Lv et al., 2023
	SMMC-7721	• Melittin	Not specified	Increases INF-γ, promote tumor cytotoxicity, enhances immune responses	Liu et al., 2016
	Huh-7	• Melittin	IC_{50}: 129.5-272.6 µg/mL	Inhibits cancer cell proliferation	El-Didamony et al., 2024
	Hepa 1-6	• Melittin	IC_{50}: 6.39 µg/mL	Induced apoptosis	Hematyar et al., 2018
Non-small cell lung cancer (NSCLC)	A549 cell (human NSCLC cell lines)	• Melittin	IC_{50}: 1 µg/mL	Inhibit migration, suppresses VEGF expression	Zhang & Chen, 2017
			IC_{50}: 2 µg/mL	Induced apoptosis, inhibits glycolysis, increase cisplatin sensitivity	Zhang et al., 2021

				Inhibits cell growth, inhibits invasion, induced TGF-β-mediated ERK pathway	Yu et al., 2021
	LLC (mouse NSCLC cell lines)	• Melittin	Not specified	Reduced M2 macrophages, increased M1/M2 ratio	Lee et al., 2017
	NCI_H460	• Melittin	IC_{50}: 3 µg/mL	Inhibit NF-kB pathway	Choi et al., 2014
	HCC1833 cell	• Melittin	IC_{50}: 0.94 to 1.49 µM for HER2-enriched cells	Inhibit proliferation, induction of autophagy cell	Wang et al., 2024
Bronchogenic carcinoma (Lung Cancer)	ChaGo-K1 (Human bronchogenic carcinoma cell line)	• Melittin	ChaGo-K1; IC_{50}: 2.5 µM	Inhibit lung cancer cell proliferation, apoptosis, inhibition of TAM	Tipgomut et al., 2018
Ovarian cancer	A2780 and A2780CR	• Melittin	A2780: Melittin IC_{50}: 5 µg/mL + Cisplatin IC_{50}: 2 µg/mL A2780CR: Melittin IC_{50}: 2 µg/mL Cisplatin: IC_{50}: 10 µg/mL	Synergistic cytotoxic effects with cisplatin, metabolic changes, induced apoptosis	Alonezi et al., 2017
	SKOV3	• Melittin	Not specified	Increases INF-γ, promote tumor cytotoxicity, enhances immune responses,	Liu et al., 2016
			IC_{50}: 80 µg/mL	Inhibits proliferation, induces cell cycle arrest (G1 phase), induced apoptosis	Su et al., 2014
Osteosarcoma	UMR-106	• Melittin	IC_{50}: 6.33 µg/mL	Inhibit angiogenesis via SDF-1α/CXCR4	Qin et al., 2016
Esophageal Squamous Cell Carcinoma (ESCC)	TE1 (human esophageal carcinoma cell line)	• Melittin	Not specified	Induced apoptosis in TE1 cells, inhibit cell proliferation in time-dependent manner	Zhou et al., 2021
Hodgkin Lymphoma (HL)	KM-H2, L-428 cell lines	• Melittin	IC_{50}: 0.93µM for KM-H2, 0.75µM for L-428	Induced apoptosis, inhibit growth, increase cisplatin sensitivity	Kreinest et al., 2020

RESULTS

A total of 35 studies met the inclusion criteria and were included in this scoping review. The selected articles investigated the anticancer activities of whole bee venom and its bioactive compounds melittin, phospholipase A2 (PLA2), apamin, adolapin, mast cell degranulating (MCD) peptide, and hyaluronidase using both *in vitro* and *in vivo* cancer models. Among 35 studies, melittin was the most extensively studied compound.

Anticancer Activity of Bee Venom and its Bioactive Compounds

The 35 included studies collectively revealed that bee venom and its bioactive components exhibit strong anticancer, antiproliferative, and pro-apoptotic activites. Most studies focused on melittin, followed by PLA2 and apamin, while fewer examined adolapin, MCD peptide, and hyaluronidase. These bioactive compounds act through multiple mechanisms involving apoptosis induction, inhibition of metastasis, and modulation of oncogenic signaling pathways.

DISCUSSION

The findings from this scoping review confirm that bee venom possesses remarkable anticancer potential, primarily attributed to its peptide and enzymatic constituents. Melittin emerged as the principal compound responsible for direct cytotoxicity through apoptosis induction and mitochondrial disruption. PLA2, apamin, adolapin, MCD peptide, and hyaluronidase complement melittin's effects by enhancing membrane permeability, modulating immune responses, and supporting intracellular signaling regulation. The multi-targeted nature of bee venom distinguishes it from conventional chemotherapeutic agents, which often act through single pathways. By influencing multiple signaling cascades such as PI3K/Akt and NF-κB, bee venom compounds not only suppress proliferation but also impede metastasis and angiogenesis.

Comparison between bioactive compounds of bee venom

Melittin demonstrates the strongest and most consistent anticancer efficacy. It triggers apoptosis, inhibits migration, and alters the tumour microenvironment (Obeidat et al., 2023). It interacts directly with the lipid bilayer of cancer cell membranes, forming pores that disrupt cellular homeostasis and lead to apoptosis (Lim et al., 2019). In addition, melittin modulates key regulatory proteins such as caspase-3, for further promoting programmed cell death. However, despite its strong efficacy, melittin's non-selective toxicity poses a major limitation such as toxicity towards both normal and cancer cells as well as other adverse effects. Studies have emphasized the need for nanoparticle-based or targeted delivery systems to minimize damage to healthy cells while maximizing therapeutic potential (Oršolić, 2011).

PLA2 plays a supporting role by enhancing melittin's cytolytic activity. It hydrolyzes phospholipids in cancer cell membranes, facilitating melittin penetration and enhancing its apoptotic effect (Yaacoub et al., 2021). Moreover, PLA2 exhibits immune-regulatory functions by activating macrophages and dendritic cells, thereby promoting antitumour immunity. Despite these benefits, its intrinsic hemolytic activity restricts its standalone use and necessities formulation refinement (Hematyar et al., 2018).

Apamin, a small neuropeptide, interferes with calcium-activated potassium channels that influence tumour cell motility. It inhibits epithelial-mesenchymal transition (EMT) a crucial step in metastasis by downregulating mesenchymal markers such as *vimentin* and *N-cadherin* (Plassay et al., 2022). Adolapin, on the other hand, exhibits strong anti-inflammatory properties that indirectly contribute to its anticancer effects by suppressing tumour-promoting cytokines like TNF-α and COX-2 (Wehbe et al., 2019). These two compounds, which exhibit lower cytotoxicity, act synergistically to reduce metastatic potential and inflammation-driven tumor progression. The MCD peptide primarily influences immune response by promoting cytokine release from mast cells, thereby enhancing immune-mediated tumour suppression (Zhang et al., 2021). Hyaluronidase, commonly referred to as the "spreading factor", enhances tissue permeability and improves the diffusion of other active peptides such as melittin and PLA2 into tumour sites (Wehbe et al., 2019). While their direct anticancer effects are limited, both components serve as important enhancers of bee venom's overall antitumour efficacy.

Comparison of Bee Venom's Anticancer Effects across Different Cancer Types

Studies on breast cancer (MCF-7 and MDA-MB-231 cells) demonstrated that melittin effectively inhibits cell proliferation and migration through suppression of the PI3K/Akt and MAPK signaling pathways (Lim et al., 2019). Apamin also contributed to reduced metastatic activity by inhibiting EMT, while hyaluronidase improved melittin diffusion within tumour tissues (Plassay et al., 2022). These findings support the potential of bee venom peptides as complementary agents in breast cancer therapy. In lung cancer (A549) models, melittin triggered apoptosis via mitochondrial disruption, while apamin inhibited EMT and cell migration (Wang et al., 2021). PLA2 enhanced these effects through synergistic membrane lysis and immune activation. Collectively, these findings suggest that combining melittin with other bee venom

components could improve therapeutic outcomes for lung carcinoma.

Melittin induced apoptosis in colon cancer cells (HT-29) through death receptor activation and NF-kB pathway inhibition, leading to reduced cell survival and inflammation (Duarte et al., 2022). PLA2 was shown to induce oxidative stress-mediated apoptosis in colon carcinoma cells. These mechanisms underline the potential of bee venom compounds as effective agents for targeting colorectal malignancies. Although limited studies are available, melittin has demonstrated potential to cross the blood-brain barrier and induce apoptosis in glioblastoma cells through activation of caspase pathways and mitochondrial disruption (Kim et al., 2015). The unique ability of melittin to penetrate neural tissue highlights its promise for future therapeutic development against brain tumors.

Collectively, the evidence presented in this review supports the hypothesis that bee venom and its bioactive constituents possess multifaceted anticancer activities through apoptosis induction, anti-metastatic effects, and immune modulation (Wehbe et al., 2019). Among these melittin remains the most potent and extensively characterized, while the synergistic effects of PLA2, apamin, adolapin, MCD peptide, and hyaluronidase enhance its overall therapeutic impact. Future research should emphasize *in vivo* models, nanocarriers formulations, and toxicity profiling to ensure selective and safe clinical applications. Bee venom therefore represents a promising, natural-based approach for the development of novel anticancer therapies.

CONCLUSION

This scoping review provides comprehensive evidence that bee venom and its bioactive compounds particularly melittin, phospholipase A2 (PLA2), apamin, adolapin, MCD peptide, and hyaluronidase exhibit significant anticancer potential through diverse biological mechanisms. Melittin emerged as the principal active compound, inducing apoptosis and suppressing tumour growth via inhibition of multiple oncogenic pathways, while other peptides contribute synergistically by modulating inflammation, enhance immune activity, and improving compound penetration into tumour tissues. Despite promising *in vivo* and clinical data highlight the need for further investigation to ensure safety, efficacy, and selectivity in therapeutic applications. Future research should focus on standardized extraction, toxicity evaluation, and the development of targeted delivery systems such as melittin-based nanoparticles. Collectively, bee venom represents a promising natural source for anticancer drug discovery and warrants continued exploration in translational cancer research.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Department of Biomedical Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM).

REFERENCES

Abd El-Wahed, A. A., Khalifa, S. A. M., Sheikh, B. Y., Farag, M. A., Saeed, A., Larik, F. A., Koca-Caliskan, U., AlAjmi, M. F., Hassan, M., Wahabi, H. A., Hegazy, M.-E. F., Algethami, A. F., Büttner, S., & El-Seedi, H. R. (2019). Bee venom composition: From chemistry to biological activity. *Studies in Natural Products Chemistry*, 459–484. https://doi.org/10.1016/b978-0-444-64181-6.00013-9

Alonezi, S., Tusiimire, J., Wallace, J., Dufton, M. J., Parkinson, J. A., Young, L. C., Clements, C. J., Park, J. K., Jeon, J. W., Ferro, V. A., & Watson, D. G. (2017). Metabolomic profiling of the synergistic effects of melittin in combination with cisplatin on ovarian cancer cells. *Metabolites*, *7*(2). https://doi.org/10.3390/metabo7020014

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I. & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer Journal Clinical*, *74*(3), 229–263. https://doi.org/10.3322/caac.21834

Ceremuga, M., Stela, M., Janik, E., Gorniak, L., Synowiec, E., Sliwinski, T., Sitarek, P., Saluk-Bijak, J., & Bijak, M. (2020). Melittin—A natural peptide from bee venom which induces apoptosis in human leukaemia cells. *Biomolecules*, *10*(2). https://doi.org/10.3390/biom10020247

Chahla, C., Rima, M., Mouawad, C., Roufayel, R., Kovacic, H., El Obeid, D., Sabatier, J. M., Luis, J., Fajloun, Z., & El-Waly, B. (2024). Effect of *Apis mellifera* syriaca bee venom on glioblastoma cancer: *In vitro* and *in vivo* studies. *Molecules*, *29*(16), 3950. https://doi.org/10.3390/molecules29163950

Choi, K., Hwang, C., Gu, S., Park, M., Kim, J., Park, J., Ahn, Y., Kim, J., Song, M., Song, H., Han, S.-B., & Hong, J. (2014). Cancer cell growth inhibitory effect of bee venom via increase of death receptor 3 expression and inactivation of NF-kappa B in NSCLC cells. *Toxins*, *6*(8), 2210–2228. https://doi.org/10.3390/toxins6082210

Cui, Z., Zhou, Z., Sun, Z., Duan, J., Liu, R., Qi, C., & Yan, C. (2024). Melittin and phospholipase A2: Promising anti-cancer candidates from bee venom. *Biomedicine &*

Pharmacotherapy, 179, 117385. https://doi.org/10.1016/j.biopha.2024.117385

Duarte, D., Falcão, S. I., El Mehdi, I., Vilas-Boas, M., & Vale, N. (2022). Honeybee venom synergistically enhances the cytotoxic effect of CNS drugs in HT-29 colon and MCF-7 breast cancer cell lines. *Pharmaceutics*, *14*(3), 511. https://doi.org/10.3390/pharmaceutics14030511

Duffy, C., Sorolla, A., Wang, E., Golden, E., Woodward, E., Davern, K., Ho, D., Johnstone, E., Pfleger, K., Redfern, A., Iyer, K. S., Baer, B., & Blancafort, P. (2020). Honeybee venom and melittin suppress growth factor receptor activation in HER2-enriched and triple-negative breast cancer. *NPJ Precision Oncology*, *4*(1). https://doi.org/10.1038/s41698-020-00129-0

El-Didamony, S. E., Kalaba, M. H., Sharaf, M. H., El-Fakharany, E. M., Osman, A., Sitohy, M., & Sitohy, B. (2024). Melittin alcalase-hydrolysate: A novel chemically characterized multifunctional bioagent; antibacterial, anti-biofilm and anticancer. *Frontiers in Microbiology*, *15*. https://doi.org/10.3389/fmicb.2024.1419917

Hematyar, M., Soleimani, M., Es-Haghi, A., & Rezaei Mokarram, A. (2018). Synergistic co-delivery of doxorubicin and melittin using functionalized magnetic nanoparticles for cancer treatment: Loading and *in vitro* release study by LC-MS/MS. *Artificial Cells, Nanomedicine, and Biotechnology*, *46*(sup3), S1226–S1235. https://doi.org/10.1080/21691401.2018.1536063

Kim, Y. W., Chatirvedi, P. K., Chun, S. N., Lee, Y. G., & Ahn, W. S. (2015). Honeybee venom possesses anticancer and antiviral effects by differential inhibition of HPV E6 and E7 expression on cervical cancer cell line. *Oncology Reports*, *33*(4), 1675–1682. https://doi.org/10.3892/or.2015.3760

Kreinest, T., Volkmer, I., & Staege, M. S. (2020). Melittin increases cisplatin sensitivity and kills KM-H2 and L-428 Hodgkin lymphoma cells. *International Journal of Molecular Sciences*, *22*(1), 343. https://doi.org/10.3390/ijms22010343

Lee, C., Bae, S. J. S., Joo, H., & Bae, H. (2017). Melittin suppresses tumor progression by regulating tumor-associated macrophages in a Lewis lung carcinoma mouse model. *Oncotarget*, *8*(33), 54951–54965. https://doi.org/10.18632/oncotarget.18627

Lim, H., Baek, S., & Jung, H. (2019). Bee venom and its peptide component melittin suppress growth and migration of melanoma cells via inhibition of PI3K/AkT/MTOR and MAPK pathways. *Molecules*, *24*(5), 929. https://doi.org/10.3390/molecules24050929

Liu, M., Wang, H., Liu, L., Wang, B., & Sun, G. (2016). Melittin-MIL-2 fusion protein as a candidate for cancer immunotherapy. *Journal of Translational Medicine*, *14*(1), 155. https://doi.org/10.1186/s12967-016-0910-0

Lv, C., Chen, J., Huang, F., Fang, F., & Li, B. (2023). Melittin inhibits the proliferation migration and invasion of HCC cells by regulating ADAMTS9-AS2 demethylation. *Toxicon*, *222*(2023), 106996–106996. https://doi.org/10.1016/j.toxicon.2022.106996

Mahmoodzadeh, A., Zarrinnahad, H., Bagheri, K. P., Moradia, A., & Shahbazzadeh, D. (2015). First report on the isolation of melittin from Iranian honey bee venom and evaluation of its toxicity on gastric cancer AGS cells. *Journal of the Chinese Medical Association*, *78*(10), 574–583. https://doi.org/10.1016/j.jcma.2015.06.008

Małek, A., Kocot, J., Mitrowska, K., Posyniak, A., & Kurzepa, J. (2022). Bee venom effect on glioblastoma cells viability and gelatinase secretion. *Frontiers in Neuroscience*, *16*(792970). https://doi.org/10.3389/fnins.2022.792970

Obeidat, M., Al-Khraisat, I. F., Jaradat, D. M. M., Ghanim, B. Y., Abdallah, Q. M., Duaa Abu Arqoub, D. A., Sabbah, D., Al-Sanabra, O. M., Arafat, T., & Qinna, N. A. (2023). Mellitin peptide quantification in seasonally collected crude bee venom and its anticancer effects on myelogenous K562 human leukaemia cell line. *BMC Complementary Medicine and Therapies*, *23*(1), 1-11. https://doi.org/10.1186/s12906-023-03897-x

Oršolić, N. (2011). Bee venom in cancer therapy. *Cancer and Metastasis Reviews*, 31(1-2), 173–194. https://doi.org/10.1007/s10555-011-9339-3

Pandey, P., Khan, F., Khan, M. A., Kumar, R., & Upadhyay, T. K. (2023). An updated review summarizing the anticancer efficacy of melittin from bee venom in several models of human cancers. *Nutrients*, 15(14), 3111. https://doi.org/10.3390/nu15143111

Plasay, M., Natzir, R., Cangara, M. H., Hardjo, M., Syahrijuita, & Soraya, G. V. (2022). Melittin-Induced cell death through p53 and 8-OHdG in breast cell cancer MCF-7. *Biomedical and Pharmacology Journal*, *15*(2), 979–983. https://biomedpharmajournal.org/vol15no2/melittin-induced-cell-death-through-p53-and-8-ohdg-in-

breast-cell-cancer-mcf-7/

PRISMA. (2020). PRISMA 2020 Flow Diagram. *PRISMA*. https://www.prisma-statement.org/prisma-2020-flow-diagram

Qin, G., Chem, Y., Li, H., Xu, S., Li, Y., Sun, J., Rao, W., Chen, C., Du, M., He, K., & Ye, Y. (2016). Melittin inhibits tumor angiogenesis modulated by endothelial progenitor cells associated with the SDF-1α/CXCR4 signaling pathway in a UMR-106 osteosarcoma xenograft mouse model. *Molecular Medicine Reports*, *14*(1), 57–68. https://doi.org/10.3892/mmr.2016.5215

Sahsuvar, S., Guner, R., Gök, Ö., & Can, O. (2023). Development and pharmaceutical investigation of novel cervical cancer-targeting and redox-responsive melittin conjugates. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-45537-x

Salimian, F., Nabiuni, M., & Salehghamari, E. (2022). Melittin prevents metastasis of epidermal growth factor-induced MDA-MB-231 cells through the inhibition of the sdf-1α/cxcr4 signalling pathway. *Cell Journal*, *24*(2), 85–90. https://doi.org/10.22074/cellj.2022.7626

Shaw, P., Kumar, N., Hammerschmid, D., Privat-Maldonado, A., Dewilde, S., & Bogaerts, A. (2025). Synergistic effects of melittin and plasma treatment: A promising approach for cancer therapy. *Cancers*, *11*(8), 1109. https://doi.org/10.3390/cancers11081109

Sisakht, M., Mashkani, B., Bazi, A., Ostadi, H., Zare, M., Avval, F. Z., Sadeghnia, H. R., Mojarad, M., Nadri, M., Ghorbani, A., & Soukhtanloo, M. (2017). Bee venom induces apoptosis and suppresses matrix metaloprotease-2 expression in human glioblastoma cells. *Revista Brasileira de Farmacognosia*, *27*(3), 324–328. https://doi.org/10.1016/j.bjp.2016.11.006

Soliman, C., Eastwood, S., Truong, V. K., Ramsland, P. A., & Elbourne, A. (2019). The membrane effects of melittin on gastric and colorectal cancer. *PLOS ONE*, *14*(10), e0224028. https://doi.org/10.1371/journal.pone.0224028

Su, M., Chang, W., Cui, M., Lin, Y., Wu, S., & Xu, T. (2014). Expression and anticancer activity analysis of recombinant human uPA1–43-melittin. *International Journal of Oncology*, *46*(2), 619–626. https://doi.org/10.3892/ijo.2014.2750

Tipgomut, C., Wongprommoon, A., Takeo, E., Ittiudomrak, T., Puthong, S., & Chanchao, C. (2018). Melittin induced G1 cell cycle arrest and apoptosis in chago-k1 human bronchogenic carcinoma cells and inhibited the differentiation of THP-1 cells into tumour- associated macrophages. *Asian Pacific Journal of Cancer Prevention: APJCP*, *19*(12), 3427–3434. https://doi.org/10.31557/APJCP.2018.19.12.3427

Wang, Y., Yuan, T., He, L., Huang, J., Wilfred, N., Yang, W., Jin, M., Huang, G., & Lu, C. (2024). Melittin treatment suppressed malignant NSCLC progression through enhancing CTSB-mediated hyperautophagy. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*, *180*(2024), 117573. https://doi.org/10.1016/j.biopha.2024.117573

Wehbe, R., Frangieh, J., Rima, M., El Obeid, D., Sabatier, J.M., & Fajloun, Z. (2019). Bee venom: Overview of main compounds and bioactivities for therapeutic interests. *Molecules*, 24(16). https://doi.org/10.3390/molecules24162997

Yaacoub, C., Rifi, M., El Obeid, D., Mawlawi, H., Sabatier, J. M., Coutard, B., & Fajloun, Z. (2021). The cytotoxic effect of *Apis mellifera* venom with a synergistic potential of its two main components—melittin and pla2—on colon cancer HCT116 cell lines. *Molecules*, *26*(8), 2264. https://doi.org/10.3390/molecules26082264

Yu, R., Wang, M., Wang, M., & Han, L. (2021). Melittin suppresses growth and induces apoptosis of non-small-cell lung cancer cells via down-regulation of TGF-β-mediated ERK signal pathway. *Brazilian Journal of Medical and Biological Research*, *54*(2). https://doi.org/10.1590/1414-431x20209017

Zamani, M., Bozorg-Ghalati, F., & Mokarram, P. (2024). Melittin as an activator of the autophagy and unfolded protein response pathways in colorectal HCT116 cell line. *Iranian Biomedical Journal*, *28*(1), 46–52. https://doi.org/10.61186/ibj.3993

Zarrinnahad, H., Mahmoodzadeh, A., Hamidi, M. P., Mahdavi, M., Moradi, A., Bagheri, K. P., & Shahbazzadeh, D. (2017). Apoptotic effect of melittin purified from Iranian honey bee venom on human cervical cancer Hela cell line. *International Journal of Peptide Research and Therapeutics*, *24*(4), 563–570. https://doi.org/10.1007/s10989-017-9641-1

Zhang, S., Lv, X., Li, L., Luo, Y., Xiang, H., Wang, L., & Li, Y. (2021). Melittin inhibited glycolysis and induced cell apoptosis in cisplatin resistant lung adenocarcinoma cells via TRIM8. *BIOCELL*, *45*(1), 167–175. https://doi.org/10.32604/biocell.2021.013636

Zhang, S. F., & Chen, Z. (2017). Melittin exerts an antitumor effect on non-small cell lung cancer cells. *Molecular Medicine Reports*, *16*(3), 3581–3586. https://doi.org/10.3892/mmr.2017.6970

Zhang, S., Liu, Y., Ye, Y., Wang, X. R., Lin, L. T., Xiao, L. Y., Zhou, P., Shi, G. X., & Liu, C. Z. (2018). Bee venom therapy: Potential mechanisms and therapeutic applications. *Toxicon*, 148, 64–73. https://doi.org/10.1016/j.toxicon.2018.04.012

Zheng, J., Lee, H. L., Ham, Y. W., Song, H. S., Song, M. J., & Hong, J. T. (2015). Anti-cancer effect of bee venom on colon cancer cell growth by activation of death receptors and inhibition of nuclear factor kappa B. *Oncotarget*, 6(42), 44437–44451. https://doi.org/10.18632/oncotarget.6295

Zhou, C., Ma, J., Lu, Y., Zhao, W., Xu, B., Lin, J., Ma, Y., Tian, Y., Zhang, Q., Wang, W., Yan, W., & Jiao, P. (2021). TERT promoter regulating melittin expression induces apoptosis and G_0/G_1 cell cycle arrest in esophageal carcinoma cells. *Oncology Letters*, *21*(1), 16. https://doi.org/10.3892/ol.2020.12277

ORIGINAL ARTICLE

Examining The Correlation Between Screen Time, Mental Health, and Dry Eye Syndrome Among Undergraduate Students at KAHS

Norzaris Arieana Zaini[1], Muhammad Afzam Shah Abdul Rahim[1,2*], Firdaus Yusof[1,2], Ilyanoon Zahari[1,2], Noor Haziq Saliman[3]

[1]Department of Optometry and Visual Science, Kuliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[2]Integrated Omics Research Group (IORG), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[3]Centre for Optometry Studies, Faculty of Health Sciences, Universiti Teknologi MARA (UiTM), Selangor, Malaysia

ABSTRACT

Background: Prolonged screen time among university students has been linked to dry eye syndrome (DES) and mental health challenges. This study examined the correlations between screen time, depression, and DES among undergraduate students at the Kuliyyah of Allied Health Sciences (KAHS), International Islamic University Malaysia (IIUM). **Methods:** A cross-sectional study of 90 students (aged 19–24 years) used validated instruments, the Screen Time Questionnaire (STQ), Patient Health Questionnaire-9 (PHQ-9), and Dry Eye-related Quality of Life Score (DEQS). **Results:** Median daily screen time was 4.2 hours, with moderate depressive symptoms (median PHQ-9 = 6) and DES impact (median DEQS = 21). Depression and DES showed a strong positive correlation (r = 0.508, $p < 0.001$), while screen time (STQ) did not correlate significantly with either. **Conclusion:** Depressive symptoms were strongly related to DES severity, suggesting psychological health should be considered in managing ocular surface complaints among students.

Keywords:
Screen time; mental health; dry eye syndrome; undergraduate students

INTRODUCTION

Prolonged screen time has been increasingly associated with adverse physical and mental health outcomes, including stress, anxiety, and depression (Twenge & Campbell, 2018). Among university students, excessive digital device use can lead to emotional distress, cognitive fatigue, and reduced physical activity (Rosenfield et al., 2011). Moreover, disturbed sleep patterns and social isolation contribute to poor psychological well-being (Rosenfield et al., 2011; Nakshine et al., 2022)

Beyond mental health implications, extended screen exposure is a well-established risk factor for dry eye syndrome (DES), a multifactorial ocular surface disorder characterised by tear film instability and discomfort (Sheppard & Wolffsohn, 2018). Digital eye strain, a common condition among heavy screen users, often coexists with DES symptoms (Coles-Brennan et al., 2018). University students are particularly vulnerable because their learning activities heavily depend on digital platforms.

Given the increased digital dependence in higher education, there is a pressing need to understand how screen time affects both mental and ocular health. This study aimed to investigate the relationships between screen time, depressive symptoms, and DES among KAHS undergraduate students at IIUM. The findings aim to provide evidence to guide preventive and management strategies for students in digital-intensive environments.

MATERIALS AND METHODS

Study Design

This cross-sectional study adhered to the Declaration of Helsinki and received ethical approval from the IIUM Research Ethics Committee (IREC 2023-KAHS/DOVS3). It

* *Corresponding author.*
E-mail address: afzamshah@iium.edu.my

investigated the relationship between screen time, mental health, and DES among undergraduate students at KAHS, IIUM. Data was collected at the IIUM Optometry Clinic to ensure standardised assessments.

The sample size was determined using GPower (version 3.1.9.2) (Faul et al., 2009), with an estimated effect size derived from prior DES prevalence studies ($\alpha = 0.05$, power = 0.80). The analysis indicated that this study required 84 participants; to account for attrition, 90 students (13 males, 77 females) were enrolled. All participants provided written informed consent after a comprehensive study briefing.

The participant cohort was recruited through convenience sampling (Golzar et al., 2022), a strategy deemed appropriate given the study's exploratory and correlational design, the accessibility of the target KAHS student population, and the practical constraints of time and resources. This methodological alignment is consistent with precedents in similar health science research (Etikan, Musa, & Alkassim, 2016). To promote sample homogeneity and mitigate confounding variables, stringent inclusion and exclusion criteria were implemented. A recognised limitation of this sampling method is its inherent vulnerability to selection bias, as volunteers may systematically differ from non-participants in characteristics such as motivation, screen-time behaviours, or health awareness. Therefore, the findings should be interpreted with caution, and their generalisability is likely confined to student populations with comparable demographics. Inclusion criteria mandated that participants be KAHS undergraduates aged 19–24 years in good systemic health, with no history of refractive surgery and visual acuity of 6/6 or better. Exclusion criteria eliminated individuals outside the specified age range, non-KAHS students, those with active systemic or ocular conditions, users of medications affecting ocular or mental health (e.g., antihistamines or antidepressants), and individuals with a history of refractive surgery or visual acuity worse than 6/6. The age restriction was controlled for developmental variability in DES susceptibility, while excluding post-refractive surgery patients minimised confounding effects on tear film stability. Visual acuity requirements ensured participants could engage with digital screens without visual-related strain, thereby isolating screen time effects more precisely.

Data Collection

All 90 participants were informed about the study and provided written consent prior to their participation. They were then asked a series of questions to gather background information and ensure they met the inclusion criteria. Objective refraction using dry retinoscopy and subjective refraction was performed to determine each participant's refractive error. Data collection employed standardised instruments to ensure reliability and validity. These included the 18-item Screen Time Questionnaire (STQ) (Carson et al., 2016; Vizcaino et al., 2019), the Patient Health Questionnaire-9 (PHQ-9) (Kroenke, Spitzer, & Williams, 2001), and the Dry Eye-Related Quality of Life Questionnaire (DEQS) (Uchino et al., 2013).

The 18-item STQ was selected to evaluate screen use across various digital devices, including smartphones, tablets, computers, and TVs. It captures both duration and context of use (academic, recreational, or social), considering weekday and weekend patterns. The STQ has demonstrated strong psychometric properties, including construct validity, test-retest reliability, and internal consistency (Carson et al., 2016). It has been validated among children, adolescents, and adults, making it suitable for use across all age groups. The questionnaire's multidimensional design was used to identify trends linked to visual symptoms, sleep disturbances, and overall well-being. Test-retest reliability has been demonstrated with ICCs ranging from 0.67 to 0.88 (Carson et al., 2016). The STQ improves on older tools by including device-specific and contextual data, enhancing the ecological validity of digital behaviour measurement (Carson et al., 2016). Using a validated tool like the STQ ensures systematic and comparable data collection. Its relevance to contemporary screen habits makes it appropriate for studies examining the health impact of digital behaviour.

The PHQ-9 was selected for this study due to its strong psychometric properties, including high validity and reliability in assessing depressive symptoms (Kroenke, Spitzer, & Williams, 2001). Developed as part of the PRIME-MD diagnostic tool, it includes nine items reflecting DSM-IV criteria for major depressive disorder, scored based on symptom frequency over the past two weeks. The PHQ-9 has demonstrated excellent internal consistency (Cronbach's alpha 0.86–0.89) and strong test-retest reliability. Sensitivity and specificity exceed 85% at the standard cutoff score of 10, supporting the clinical and research utility of this measure. Its brief, self-administered format makes it practical for mental health screening in diverse populations. The PHQ-9's minimal burden, strong psychometric performance, and sensitivity to meaningful changes make it suitable for detecting at-risk individuals and monitoring mental health status in academic or clinical settings.

The DEQS questionnaire was chosen for this study due to its established reliability, validity, and sensitivity in

evaluating the impact of dry eye symptoms on quality of life (Uchino et al., 2013; Sakane et al., 2013). Developed specifically for dry eye patients, the DEQS assesses both symptom frequency and the resulting functional and emotional burden using a 15-item scale. Validation studies have demonstrated strong internal consistency (Cronbach's alpha = 0.88) and high test-retest reliability (ICC = 0.91), as well as good convergent validity with other dry eye symptom questionnaires and clinical indicators (Uchino et al., 2013). The DEQS also avoids significant floor and ceiling effects, supporting its suitability across varying levels of dry eye severity. The questionnaire's design allows for a single summary score ranging from 0 (no impact) to 100 (maximum impact), enabling clinicians and researchers to quantify subjective symptom burden with precision. Its ease of use, brief format, and robust psychometric profile make the DEQS particularly valuable for identifying individuals at risk and evaluating the effects of treatment or environmental factors on ocular surface-related quality of life.

Statistical Analysis

Data were analysed using SPSS software (Version 29; IBM Corp., Chicago, IL, USA). The Shapiro-Wilk test revealed non-normal distributions for all key variables, consistent with Ghasemi (2012) recommendations for small sample sizes, where the Shapiro-Wilk test is more powerful than the Kolmogorov-Smirnov test for detecting non-normality in small-to-moderate samples ($n < 200$), making it the preferred choice when sample sizes are below 100. Consequently, as suggested by Puth et al. (2015), Spearman's rank-order correlation (ρ) was employed to examine three primary relationships: (i) screen time (STQ) versus mental health (PHQ-9), (ii) screen time (STQ) versus dry eye syndrome (DEQS), and (iii) mental health (PHQ-9) versus dry eye syndrome (DEQS). Effect sizes were interpreted using Cohen (1988) criteria (0.1-0.3 = weak, 0.3-0.5 = moderate, >0.5 = strong), with statistical significance set at $p < 0.05$.

RESULTS

Demographic Information

A total of 90 participants were enrolled (females: n = 77; males: n = 13). The median age of participants was 21.0 years (IQR: 20.0–21.0; range: 19–24). When analyzed by gender, males had a median age of 20.0 years (IQR: 19.0–21.0), and females had a median age of 21.0 years (IQR: 20.0–21.0). Refractive error comparisons between right and left eyes also revealed no statistically significant differences (Wilcoxon signed-rank test, $p > 0.05$), thus, only right eye data were used for further analysis.

Table 1 provides the median (Q1, Q3) values for key parameters assessed in the study, both for the total sample (n = 90) and separately by gender (female: n = 77; male: n = 13).

Table 1: Median and interquartile range (IQR) of the investigated parameters for the participants and by gender

Parameters	Median (Q1, Q3)		
	Male (n=13)	Female (n=77)	Total (n=90)
Degree of Myopia, Spherical Equivalent (D)	−1.75 (−3.31, −0.18)	−1.38 (−2.50, −0.25)	−1.50 (−2.75, −0.25)
STQ (Total Score)	4.15 (3.25, 5.05)	4.25 (3.42, 5.07)	4.22 (3.37, 5.08)
PHQ-9 (Total Score)	5.50 (3.38, 7.63)	6.00 (3.25, 8.75)	6.00 (3.38, 8.63)
DEQS (%)	22.10 (14.45, 29.75)	20.84 (6.34, 35.34)	20.84 (6.47, 35.22)

Relationship between STQ and PHQ-9

Analysis revealed a weak, non-statistically significant positive correlation between STQ and PHQ-9 scores, r (88) = 0.116, p = 0.276, (Figure 1).

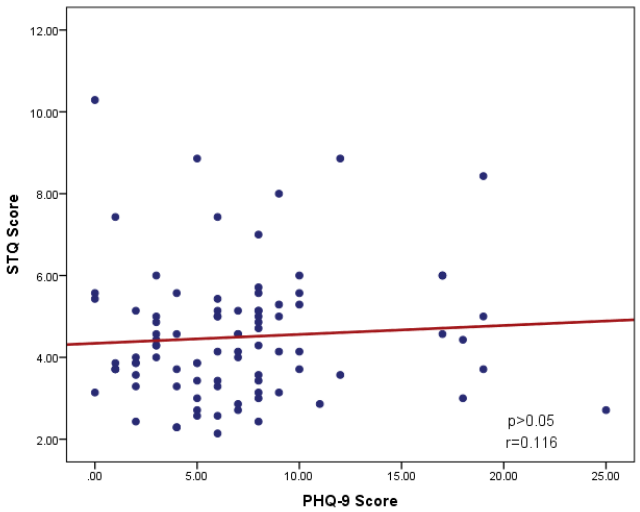


Figure 1: Scatter plot showing the correlation between STQ Score and PHQ-9 Score

Relationship between STQ and DEQS

There was a weak, non-statistically significant, positive correlation between STQ and DEQS scores, r (88) = 0.182, p = 0.085, (Figure 2).

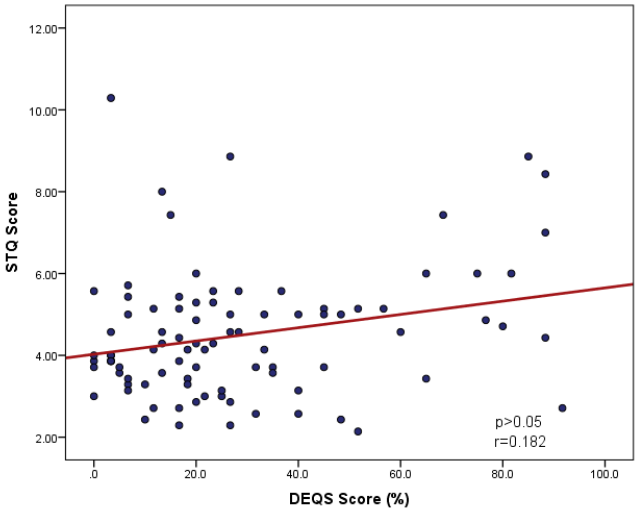


Figure 2: Scatter plot showing the correlation between STQ Score and DEQS Score (%)

Relationship between PHQ-9 and DEQS

A strong, statistically significant positive correlation was found between PHQ-9 and DEQS scores (r = 0.508, p < 0.001) (Figure 3). This indicates that higher depression scores were strongly associated with more severe dry eye symptoms. The effect size (r = 0.508) represents a strong and meaningful relationship, suggesting that depressive symptoms substantially contribute to the perceived impact of DES on quality of life, rather than this association being merely statistically significant (Cohen, 1988).

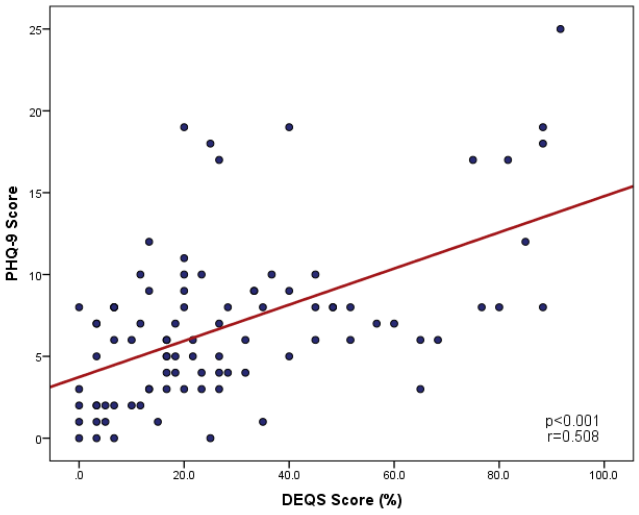


Figure 3: Scatter plot showing the correlation between PHQ-9 Score and DEQS Score (%)

DISCUSSION

This study examined the correlations between screen time (STQ), mental health (PHQ-9), and dry eye syndrome (DEQS) among undergraduate students at KAHS. The findings revealed several important relationships that contribute to our understanding of how digital device usage, psychological well-being, and ocular health interact in academic populations. Results showed no statistically significant correlation between the STQ and PHQ-9 (Figure 1) or between the STQ and DEQS (Figure 2). However, a statistically significant correlation was found between the PHQ-9 and DEQS (Figure 3). These results provided valuable insights while also raising questions about the complex interactions between these factors.

The absence of a significant correlation between screen time (STQ) and depressive symptoms (PHQ-9) contrasts with findings from previous studies that reported stronger associations between screen exposure and poor psychological well-being. For instance, Tang et al. (2021) and Wu et al., (2016) conducted a systematic review of longitudinal studies and found that higher recreational screen use was consistently linked with increased risks of depression among young people. Similarly, Twenge and Campbell (2018) observed that excessive screen time predicted lower psychological well-being in adolescents from population-based data. The divergence between those findings and the present study may be explained by contextual differences in screen-use behaviour. The current participants, who were health science students, predominantly used screens for academic rather than leisure purposes, which may involve more structured and purposeful engagement. Academic screen use has been associated with less emotional burden compared with passive or entertainment-based use (Lissak, 2018). Therefore, the educational nature of screen activity and heightened awareness of eye and mental health among KAHS students may have mitigated the adverse effects reported in general populations.

Similarly, the weak correlation between screen time and DES symptoms was somewhat surprising given the established literature on digital eye strain. Several factors may account for this finding. First, all participants had normal or corrected-to-normal visual acuity (6/6 or better), which may have provided some protection against screen-related ocular strain. Second, the average daily reported screen time (median, 4.22 hours; Table 1) was below the threshold associated with more severe DES symptoms reported in other studies. For example, Li et al. (2012) found a notable increase in DED symptoms in adults using digital devices for 6 hours or more daily. It is possible that health science students, being more aware of ocular

health, may unconsciously employ protective behaviours such as more frequent blinking or taking regular breaks from screens. Environmental factors not measured in this study, such as room lighting and humidity levels, might also have influenced these results.

The most significant finding was the strong positive correlation between depressive symptoms and DES (Figure 3). This aligns with growing evidence suggesting bidirectional relationships between mental health and ocular surface disorders. From a physiological perspective, stress and depression may affect tear production and ocular surface health through neurohormonal pathways (Zhang et al., 2019). Psychologically, the discomfort and visual disturbances associated with DES may contribute to increased stress and reduced quality of life, potentially exacerbating depressive symptoms (Ayaki et al., 2015). The moderate impact of DES on quality of life, as observed in our study with a median DEQS score of 20.84 (Table 1), supports this interpretation. These findings highlight the importance of considering psychological factors in the assessment and management of DES, particularly in student populations where academic pressures may compound these effects.

STUDY LIMITATIONS

This study had several limitations that should be carefully considered when interpreting the findings and applying them to broader contexts. First, the cross-sectional design limits the ability to establish causality between screen time, mental health, and dry eye syndrome. While associations can be observed, the temporal direction of these relationships remains unclear. Longitudinal studies or randomised controlled trials would be more appropriate for identifying temporal sequences and causal links (Levin, 2007).

Second, the use of self-reported instruments, namely the STQ, PHQ-9, and DEQS, introduced potential sources of information bias, such as recall bias and social desirability bias. Participants may have under- or overestimated their screen usage or symptoms either intentionally or unintentionally (Althubaiti, 2016). In addition, screen time data were not verified with objective digital tracking tools, which could have improved accuracy. Environmental and contextual factors, such as lighting conditions, air quality, and screen ergonomics, were also not assessed, although they are known to influence both mental health and ocular discomfort.

Another limitation is the relatively homogeneous demographic composition of the study sample, which consisted of undergraduate students from a single institution. Although participants came from a range of health science programs including optometry, dietetics, physiotherapy, radiology, audiology, speech therapy, and biomedical sciences, they shared somewhat similar academic pressures, age range, and institutional environment. This limited variability may reduce the external validity and generalizability of the findings to broader populations, such as those from different educational backgrounds, age groups, or individuals in varied occupational settings with diverse screen time behaviours and baseline health statuses (Bornstein et al., 2013).

Another demographic limitation was the gender imbalance among participants, with only 13 males and 77 females. Although gender was not a variable investigated in this study, this disproportion may affect the representativeness of the findings and should be considered in future research designs.

Additionally, this study relied solely on questionnaire-based screening tools without incorporating actual, measurable clinical diagnostics. Although the PHQ-9 and DEQS are widely validated and used in research and screening settings, they do not replace structured clinical interviews or comprehensive eye examinations, which can yield a more accurate diagnosis of depression or dry eye disease (Kroenke et al., 2001; Uchino et al., 2013; Okumura et al., 2020). The absence of such clinical evaluations limits the precision of symptom classification.

RECOMMENDATIONS FOR FUTURE RESEARCH

Future research should consider adopting a longitudinal design to establish causal relationships between screen time, mental health, and dry eye syndrome. Tracking participants over time would provide valuable insight into the direction and persistence of these associations. Expanding the study population to include students from multiple institutions and diverse academic disciplines would also enhance the generalisability of the findings. In addition, future studies could address current methodological limitations through practical improvements. The use of objective digital tracking applications on smartphones or computers could provide more accurate measures of screen exposure, thereby reducing recall bias that is often associated with self-reported questionnaires (Lin et al., 2016). Incorporating clinical eye examinations, such as Schirmer's test, tear breakup time, and ocular surface staining, would allow researchers to confirm questionnaire-based findings and strengthen the diagnostic validity of dry eye assessments (Stapleton et al., 2017). Finally, recruiting larger and more diverse samples would increase statistical power and allow

meaningful subgroup analyses, thereby advancing understanding of how digital behaviour influences ocular and psychological health.

CONCLUSION

In conclusion, depressive symptoms were strongly associated with the severity of dry eye symptoms, whereas screen time alone did not show a significant relationship with either variable. These findings highlight the close interaction between psychological and ocular health among students. Eye care professionals should consider screening for stress or depression when treating students with dry eye complaints, as addressing underlying mental health factors may improve overall management outcomes. Educators can play a preventive role by encouraging students to take regular screen breaks and practise good visual ergonomics during prolonged digital tasks. Together, these approaches can help protect both ocular comfort and mental well-being in academic environments that demand extensive screen use.

ACKNOWLEDGEMENT

This research was not funded by any grant. The authors would like to thank all participants who took part in this study, as well as the academic staff of the institution for their support and assistance during data collection.

REFERENCES

Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. Journal of Multidisciplinary Healthcare, 211–211. https://doi.org/10.2147/jmdh.s104807

Ayaki, M., Kawashima, M., Negishi, K., & Tsubota, K. (2015). High prevalence of sleep and mood disorders in dry eye patients: survey of 1,000 eye clinic visitors. Neuropsychiatric Disease and Treatment, 889. https://doi.org/10.2147/ndt.s81515

Bornstein, M. H., Jager, J., & Putnick, D. L. (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. Developmental Review, 33(4), 357–370. https://doi.org/10.1016/j.dr.2013.08.003

Carson, V., Hunter, S., Kuzik, N., Gray, C.E., Poitras, V.J., Chaput, J.P., Saunders, T.J., Katzmarzyk, P.T., Okely, A.D., Connor Gorber, S., Kho, M.E., Sampson, M., Lee, H. & Tremblay, M.S. (2016). Systematic review of sedentary behaviour and health indicators in school-aged children and youth: an update. Applied Physiology, Nutrition and Metabolism, 41(6 Suppl 3):S240-65. https://doi.org/10.1139/apnm-2015-0630

Cohen, J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.

Coles-Brennan, C., Sulley, A., & Young, G. (2018). Management of digital eye strain. Clinical and Experimental Optometry, 102(1), 18–29. https://doi.org/10.1111/cxo.12798

Etikan, I., Musa, S.A. & Alkassim, R.S. (2016). Comparison of convenience sampling and purposive sampling. American Journal of Theoretical and Applied Statistics, 5(1):1–4 https://doi.org/10.11648/j.ajtas.20160501.11

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. International Journal of Endocrinology and Metabolism, 10(2), 486–489. https://doi.org/10.5812/ijem.3505

Golzar, J., Tajik, O., & Noor, S. (2022). Convenience sampling. ResearchGate. https://doi.org/10.22034/ijels.2022.162981

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine, 16(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Levin, K. A. (2007). Study design VII. Randomised controlled trials. Evidence-Based Dentistry, 8(1), 22–23. https://doi.org/10.1038/sj.ebd.6400473

Li, M., Gong, L., Chapin, W. J., & Zhu, M. (2012). Assessment of vision-related quality of life in dry eye patients. Investigative Ophthalmology and Visual Science, 53(9), 5722–5722. https://doi.org/10.1167/iovs.11-9094

Lin, L.Y., Sidani, J.E., Shensa, A., Radovic, A., Miller, E., Colditz, J.B., Hoffman, B.L., Giles, L.M. & Primack, B.A. (2016). Association between social media use and

depression among US young adults. Depression and Anxiety, Apr;33(4):323-31. https://doi.org/10.1002/da.22466

Lissak, G. (2018). Adverse physiological and psychological effects of screen time on children and adolescents: Literature review and case study. Environmental Research, 164(1), 149–157. https://doi.org/10.1016/j.envres.2018.01.015

Nakshine, V. S., Thute, P., Khatib, M. N., & Sarkar, B. (2022). Increased screen time as a cause of declining physical, psychological health, and sleep patterns: A literary review. Curēus, 8;14(10):e30051. https://doi.org/10.7759/cureus.30051

Okumura, Y., Inomata, T., Iwata, N., Sung, J., Fujimoto, K., Fujio, K., Inomata, A.M., Miura, M., Akasaki, Y. & Murakami, A. (2020). A review of dry eye questionnaires: Measuring patient-reported outcomes and health-related quality of life. Diagnostics, 10(8), 559–559. https://doi.org/10.3390/diagnostics10080559

Rosenfield, M. (2011). Computer vision syndrome: a review of ocular causes and potential treatments. Ophthalmic and Physiological Optics, 31(5), 502–515. https://doi.org/10.1111/j.1475-1313.2011.00834.x

Sakane, Y., Yamaguchi, M., Yokoi, N., Uchino, M., Dogru, M., Oishi, T., Ohashi, Y., & Ohashi, Y. (2013). Development and validation of the dry eye–related quality-of-life score questionnaire. JAMA Ophthalmology, 131(10), 1331–1331. https://doi.org/10.1001/jamaophthalmol.2013.4503

Sheppard, A. L., & Wolffsohn, J. S. (2018). Digital eye strain: prevalence, measurement and amelioration. BMJ Open Ophthalmology, 3(1), e000146. https://doi.org/10.1136/bmjophth-2018-000146

Stapleton, F., Alves, M., Bunya, V.Y., Jalbert, I., Lekhanont, K., Malet, F., Na, K.S., Schaumberg, D., Uchino, M., Vehof, J., Viso, E., Vitale, S. & Jones, L. (2017). TFOS DEWS II epidemiology report. The Ocular Surface, Jul;15(3):334-365. https://doi.org/10.1016/j.jtos.2017.05.003

Tang, S., Werner-Seidler, A., Torok, M., Mackinnon, A. J., & Christensen, H. (2021). The relationship between screen time and mental health in young people: A systematic review of longitudinal studies. Clinical Psychology Review, 86(86), 102021. https://doi.org/10.1016/j.cpr.2021.102021

Twenge, J. M., & W. Campbell, W.K. (2018). Associations between screen time and lower psychological well-being among children and adolescents: Evidence from a population-based study. Preventive Medicine Reports, 12, 271–283. https://doi.org/10.1016/j.pmedr.2018.10.003

Uchino, M., Yokoi, N., Uchino, Y., Dogru, M., Kawashima, M., Komuro, A., Sonomura, Y., Kato, H., Kinoshita, S., Schaumberg, D. A., & Tsubota, K. (2013). Prevalence of dry eye disease and its risk factors in visual display terminal users: The Osaka study. American Journal of Ophthalmology, 156(4), 759-766.e1. https://doi.org/10.1016/j.ajo.2013.05.040

Vizcaino, M., Buman, M., DesRoches, C.T., Wharton, C. (2019). Reliability of a new measure to assess modern screen time in adults. BMC Public Health, 19(1). https://doi.org/10.1186/s12889-019-7745-6

Wu, X., Tao, S., Zhang, S., Zhang, Y., Chen, K., Yang, Y., Hao, J., & Tao, F. (2016). Impact of screen time on mental health problems progression in youth: A 1-year follow-up study. BMJ Open, 6(11), e011533–e011533. https://doi.org/10.1136/bmjopen-2016-011533

Zhang, X., Yin, Y., Yue, L., & Gong, L. (2019). Selective serotonin reuptake inhibitors aggravate depression-associated dry eye via activating the NF-κB pathway. Investigative Opthalmology & Visual Science, 60(1), 407. https://doi.org/10.1167/iovs.18-25572

ORIGINAL ARTICLE

The Effects of Water Consumption on Biometry in Myopic Young Adults

Amirah Irdina As'ari[1], Nureen Syaza Che Mohd Zain[1], Muhammad Afzam Shah Abdul Rahim[1, 2], Mizhanim Mohamad Shahimin[3], Sabrina Subri[4], Firdaus Yusof [1,2*]

[1]Department of Optometry and Visual Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Kuantan, Pahang Darul Makmur, Malaysia.
[2]Integrated Omics Research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Kuantan, Pahang Darul Makmur, Malaysia.
[3]Optometry & Vision Science Programme, Centre for Community Health Studies (ReaCH), Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur Campus, Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia.
[4]Center of Optometry, Faculty of Health Sciences, Universiti Teknologi MARA, UiTM Cawangan Selangor, 42300 Puncak Alam, Selangor, Malaysia.

Keywords:
Myopia; axial length; biometry; hydration; water drinking test

ABSTRACT

Background: Myopia, characterized by excessive axial elongation and altered scleral biomechanics, is associated with an increased risk of ocular morbidity. Although axial length (AL) measurement is crucial for monitoring myopia progression, its stability under physiological stressors such as water ingestion remains uncertain. The water drinking test (WDT) transiently elevates intraocular pressure (IOP) and could potentially deform ocular structures, particularly in eyes with reduced scleral compliance. This study investigated the acute effects of water consumption on ocular biometry; AL, central corneal thickness (CCT), anterior chamber depth (ACD), and lens thickness (LT) in myopic and emmetropic young adults. **Methods:** Thirty-two healthy participants (15 emmetropes, 17 myopes; aged 18–25 years) underwent standardized WDT involving ingestion of 100 mL of water per 10 kg of body weight. Biometric parameters were measured at baseline and at 5, 15, 30, and 45 minutes post-ingestion using Lenstar optical biometry (Haag-Streit, Switzerland). Data were analyzed using appropriate parametric and non-parametric tests, with $p < 0.05$ considered statistically significant. **Results:** AL remained stable in myopes ($F(4,64)=1.678$, $p=0.166$) but showed a modest time effect in emmetropes ($\chi^2(4)=10.354$, $p=0.035$). Myopes exhibited consistently longer AL than emmetropes across all time points ($p<0.05$). LT showed a time effect in emmetropes ($p=0.004$) without significant pairwise changes. No significant alterations were observed for CCT or ACD in either group. **Conclusion:** Acute systemic hydration did not significantly affect ocular biometry in myopic young adults, supporting the reliability of single-timepoint measurements for clinical monitoring. The persistent AL disparity between groups highlights inherent structural differences in ocular biomechanics independent of short-term hydration.

INTRODUCTION

Myopia is an escalating global public health concern, projected to affect nearly half of the world's population by 2050 (Holden et al., 2016). The condition arises primarily from excessive elongation of the eyeball, resulting in light focusing anterior to the retina and subsequent visual blur. Beyond its refractive implications, myopia predisposes individuals to sight-threatening complications such as retinal detachment, myopic maculopathy, and choroidal neovascularization (Ohno-Matsui et al., 2021). The key structural determinant of myopia is axial length (AL), which reflects the degree of ocular elongation and serves as a critical metric for progression monitoring and intervention efficacy.

Biomechanically, myopic eyes undergo profound alterations within the scleral tissue. Chronic axial elongation triggers extracellular matrix remodeling, leading to thinning, collagen fiber disorganization, and reduced mechanical stiffness (McBrien et al., 2009; Yu & Zhou, 2022). These changes compromise the sclera's ability to resist deformation from intraocular pressure (IOP) or other transient forces, thereby influencing both ocular structure and function. Furthermore, reduced choroidal blood flow in myopia may disturb scleral hydration and further weaken its biomechanical stability (Zhou et al., 2021).

The water drinking test (WDT), originally developed as a provocative test for glaucoma, provides a controlled physiological challenge that transiently elevates IOP following rapid fluid ingestion (Read & Collins, 2010). This transient IOP elevation reflects a short-term imbalance in aqueous dynamics that may deform ocular structures,

[1, 2*] *Corresponding author.*
E-mail address: yfirdaus@iium.edu.my

particularly in eyes with altered biomechanics. Recent local data by Yusof et al. (*in press*) demonstrated that systemic fluid ingestion significantly increases IOP in normal young Malaysian adults, confirming that the WDT effectively induces measurable intraocular stress under real-world conditions. This finding supports the use of WDT as a physiological model to probe the biomechanical resilience of the eye.

Prior studies have reported inconsistent outcomes regarding ocular biometry following WDT. Some observed transient axial shortening in myopic eyes, possibly due to pressure-induced compression of a thinner sclera or choroidal volume changes (Read & Collins, 2010), whereas others found axial elongation, implying heterogeneous scleral responses to fluid loading (Junghans et al., 2013). Given these contradictions, the influence of systemic hydration on ocular biometry remains unclear. Understanding whether hydration affects parameters such as AL, central corneal thickness (CCT), anterior chamber depth (ACD), and lens thickness (LT) is crucial, as such variability could confound biometry-based clinical monitoring in myopia management.

Therefore, this study aimed to investigate the short-term effects of water ingestion on ocular biometry in myopic and emmetropic young adults. By examining time-dependent changes following WDT, this study sought to determine whether myopic eyes exhibit distinct biomechanical responses to transient physiological stress, thereby assessing the reliability of biometric measurements in routine clinical and research settings.

MATERIALS AND METHODS

This cross-sectional study was conducted at the Optometry Clinic, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM) Kuantan. Ethical approval was obtained from the IIUM Research Ethics Committee (IREC) (Approval No. IREC 2024-KAHS/DOVS2). The study adhered to the principles of the Declaration of Helsinki for research involving human participants.

Participants

Healthy volunteers aged 18–25 years were recruited and categorized into two refractive groups: emmetropes (spherical equivalent +0.50 D to –0.25 D) and myopes (spherical equivalent < –0.50 D) (Morgan et al., 2010). Participants with ocular disease, prior ocular surgery, pregnancy, systemic illness, or media opacity were excluded. All participants were instructed to abstain from food and drink (including caffeine intake) for at least two hours before testing (Hamdan et al., 2024). Participants were tested at a relatively similar time to prevent diurnal effects on the IOP (Kim et al., 2016).

Sample Size Determination

The target sample size was determined a priori for a mixed (between–within) repeated-measures ANOVA with two groups (myopic vs emmetropic) and five time points (baseline, 5, 15, 30, 45 minutes). Assuming a small-to-medium effect size ($f = 0.25$), $\alpha = 0.05$, power $(1–\beta) = 0.80$, an estimated correlation among repeated measures $r = 0.60$, and nonsphericity correction $\varepsilon = 0.70$, the required total sample size was $N = 28$ (≈ 14 per group). Allowing for attrition, the target enrollment was 30–32 participants. The final sample comprised 32 participants (emmetropes $n = 15$; myopes $n = 17$), satisfying this requirement and providing 80% power to detect $f \geq 0.25$ effects on axial length. Sample size estimation was performed using GPower v3.1 (Faul et al., 2007).

Water Drinking Test Protocol

Following baseline measurements (T_0), each participant consumed 100 mL of water per 10 kg of body weight within five minutes, following the standard water drinking test (WDT) protocol (Raman & Clement, 2020; Yusof et al., 2024). Ocular biometry and systemic parameters were then recorded at 5, 15, 30, and 45 minutes post-ingestion (T_1–T_4).

Ocular and Systemic Measurements

Ocular biometry, including axial length (AL), central corneal thickness (CCT), anterior chamber depth (ACD), and lens thickness (LT), was measured using the Lenstar LS 900 optical biometer (Haag-Streit, Switzerland). Systemic parameters (systolic (SBP) and diastolic (DBP) blood pressure) were measured concurrently using an OMRON HEM-907 digital sphygmomanometer (Omron, Japan), and oxygen saturation (SpO_2) and perfusion index (PI) were recorded using a fingertip pulse oximeter.

Data Analysis

All data were analyzed using IBM SPSS Statistics version 20.0 (IBM Corp., Armonk, NY). The normality of data distribution was assessed using the Shapiro-Wilk test to determine the appropriate statistical approach. For normally distributed data, repeated-measures analysis of variance (ANOVA) was employed to evaluate within-group changes across time points, with Greenhouse-Geisser corrections applied whenever the assumption of sphericity was violated, as determined by Mauchly's test. Non-normally distributed repeated measures were analyzed using the Friedman test. Comparisons between the myopic and emmetropic groups at each time point were conducted using either the independent-samples t-test for normally distributed variables or the Mann-Whitney U test

for non-parametric data. To examine the combined effects of time and refractive group, a mixed (between-within) ANOVA model was applied. Statistical significance was set at $p < 0.05$ for all analyses, and relevant effect sizes were reported where appropriate.

RESULTS

Demographic and Baseline Characteristics

A total of 32 participants were included, 15 emmetropes (11 females, 4 males) and 17 myopes (all females). A significant age difference was observed between the groups (U = 77, z = –2.089, p = 0.037), with emmetropes (median = 23 years) being slightly older than myopes (median = 22 years). No significant differences were found in weight (U = 94.50, p = 0.210) or baseline pulse rate (t(30) = 0.124, p = 0.902). As expected, participants' refractive errors in spherical equivalent (SE) differed significantly between groups (U < 0.001, p < 0.001), with myopes showing higher refractive error (median = -6.00 D) compared to emmetropes (median = –0.25 D). Baseline axial length (AL) was also significantly longer in myopes (U = 27, p < 0.001). The demographic and baseline characteristics are summarized in Table 1.

Table 1: Demographic and baseline ocular characteristics of emmetropic and myopic participants. Data are presented as mean ± standard deviation (SD) for normally distributed variables and as median (interquartile range, IQR) for non-normal data (#). *P*-values represent between-group comparisons (emmetropic vs myopic), either using the t-test or the Mann-Whitney U test. Significant differences at $p < 0.05$ are indicated in bold.

Variable	Emmetropic (n=15, 11 females, 4 males)	Myopic (n=17, 17 females)	Test Statistic	p
Age (years)#	23 (1)	22 (2)	*U* = 77	**0.037**
Weight (kg)#	65 (30)	50 (25)	*U* = 94.50	0.210
Spherical Equivalent, SE (D)#	-0.25 (0.25)	-6.00 (3.75)	*U* < 0.001	**<0.001**
Pulse Rate (BPM)	76.97 ± 6.21	76.59 ± 10.45	t (30) = 0.124	0.902
Baseline AL (mm)#	23.86 (0.69)	25.91 (1.81)	*U* = 27	**<0.001**

Axial Length (AL)

For emmetropic eyes, AL data were non-normally distributed (Shapiro-Wilk $p < 0.05$). A Friedman test revealed a significant difference in AL across time points ($\chi^2(4) = 10.354$, $p = 0.035$). Post-hoc pairwise comparisons with a Bonferroni adjustment revealed a significant decrease in AL from T_1 to T_3 ($p = 0.032$). In contrast, AL data in myopic eyes were normally distributed ($p > 0.05$), and repeated-measures ANOVA showed no significant variation over time ($F(4,64) = 1.678$, $p = 0.166$). Between-group comparisons at each time point showed consistently greater AL in myopes than emmetropes ($p < 0.05$ at all intervals). Figure 1 illustrates AL trends across time points in both groups.

Central Corneal Thickness (CCT)

Mauchly's test indicated violation of sphericity in both emmetropic (W = 0.150, p = 0.006) and myopic groups

(W = 0.206, p = 0.007); therefore, Greenhouse-Geisser corrections were applied. Repeated-measures ANOVA revealed no significant temporal change in CCT for emmetropes ($F(2.105,29.475) = 0.463$, $p = 0.644$, partial $\eta^2 = 0.032$) or myopes ($F(2.211,35.370) = 1.094$, $p = 0.351$, partial $\eta^2 = 0.064$).

Anterior Chamber Depth (ACD)

ACD data were normally distributed except at one time point (T_2, $p = 0.023$). Using repeated-measures ANOVA with Greenhouse-Geisser correction (emmetropes: W = 0.037, p < 0.005; myopes: W = 0.223, p = 0.011), no significant effect of time was found for either group (emmetropes: $F(1.661,23.248) = 2.569$, $p = 0.106$, partial $\eta^2 = 0.155$; myopes: $F(2.345,37.523) = 0.819$, $p = 0.456$, partial $\eta^2 = 0.049$).

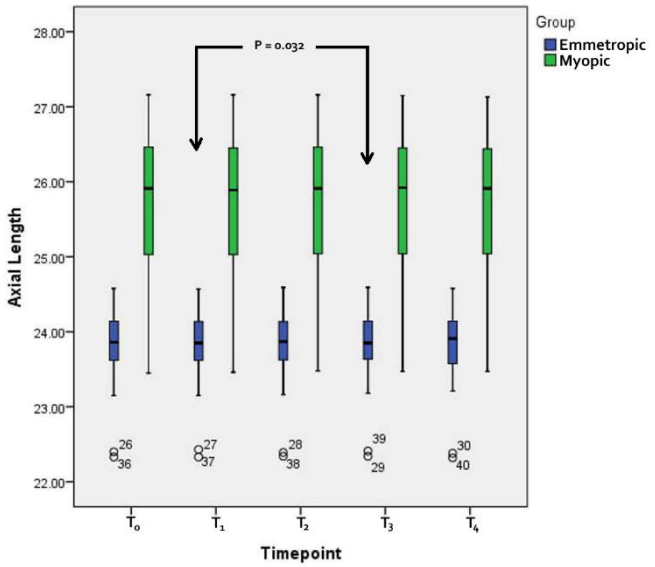


Figure 1: Distribution of axial length (AL) measurements in emmetropic and myopic eyes at baseline (T_0) and four subsequent timepoints (T_1–T_4) following water ingestion. Boxes represent the interquartile range (IQR), the horizontal line indicates the median, whiskers denote 1.5×IQR, and circles represent outliers. The Bonferroni post hoc test indicates a significantly greater AL at T_1 than at T_3.

Lens thickness (LT)

LT data met normality assumptions across all time points. In emmetropes, sphericity was violated (χ^2 (9) = 43.361, p < 0.001); after Greenhouse–Geisser correction, a significant time effect was observed (F (1.870,26.184) = 7.319, p = 0.004, partial η^2 = 0.343). Post-hoc pairwise comparisons with a Bonferroni adjustment revealed a significant decrease in LT from T_0 to T_2 (p = 0.004), from T_3 to T_2 (p = .001), and from T_4 to T_2 (p=0.014). In myopes, no significant temporal changes were detected (F (2.669,42.711) = 0.372, p = 0.751, partial η^2 = 0.023).

Systemic Parameters

Systolic Blood Pressure (SBP) showed a significant overall effect of time (emmetropes: F(4,56) = 3.248, p = 0.017, η^2 = 0.169; myopes: F(4,56) = 2.566, p = 0.048, η^2 = 0.115), but no significant time × group interaction (p = 0.088). Diastolic Blood Pressure (DBP) showed no significant changes over time (emmetropes: $\chi^2(4)$ = 3.891, p = 0.421; myopes: F(1.81,29.03) = 2.027, p = 0.153).

Oxygen Saturation (SpO_2) remained stable across all time points in both groups (emmetropes: $\chi^2(4)$ = 1.221, p = 0.875; myopes: $\chi^2(4)$ = 7.300, p = 0.121). Perfusion Index (PI) showed significant decreases across time in both groups (emmetropes: p = 0.003; myopes: p = 0.002–0.001), but no between-group difference (p > 0.05). Table 2 summarizes the biometric and systemic data across time intervals.

Table 2: Mean ± standard deviation (or median for non-normal data) of ocular and systemic parameters: axial length (AL), central corneal thickness (CCT), anterior chamber depth (ACD), lens thickness (LT), systolic blood pressure (SBP), diastolic blood pressure (DBP), oxygen saturation (SPO_2), and perfusion index (PI), measured at five time intervals (T_0 = baseline; T_1 = 5 min; T_2 = 15 min; T_3 = 30 min; T_4 = 45 min) following water ingestion in emmetropic and myopic participants. *p* values represent within-group comparisons across time points. Data are expressed as mean ± standard deviation (SD) for normally distributed variables and as median (interquartile range, IQR) for non-normal data.

Group	Parameter	Time Interval					p*
		T_0	T_1	T_2	T_3	T_4	
Emmetropic (n = 15)	**AL (mm)‡**	23.86 (0.69)	23.85 (0.69)	23.87 (0.69)	23.81 (0.68)	23.91 (0.71)	**0.035†**
	CCT (mm)	555.13 ± 31.30	554.27 ± 32.88	555.33 ± 32.99	555.20 ± 33.05	555.13 ± 32.99	0.644
	ACD (mm)	3.20 ± 0.15	3.20 ± 0.15	3.21 ± 0.15	3.20 ± 0.16	3.19 ± 0.17	0.106
	LT (mm)	3.48 ± 0.18	3.47 ± 0.17	3.45 ± 0.19	3.48 ± 0.19	3.49 ± 0.19	**0.004†**
	SBP (mmHg)	112.2 ± 13.0	109.0 ± 11.0	107.1 ± 15.6	108.1 ± 15.0	109.33 ± 13.1	**0.017†**
	DBP (mmHg)‡	71.0(14)	71.0(12)	68.0(13)	74.0(13)	71.0(15)	0.421
	SPO2 (%)‡	98(2)	98(1)	98(0)	98(1)	98(0)	0.875
	PI (%)‡	3.4(2.7)	2.1(3.2)	3.2(3.7)	2.1(2.7)	2.3(1.4)	**0.021†**
Myopic (n = 17)	**AL**	25.64 ± 1.12	25.64 ± 1.12	25.65 ± 1.12	25.65 ± 1.12	25.64 ± 1.12	0.166
	CCT	549.35 ± 22.16	547.94 ± 21.13	548.82 ± 20.29	549.24 ± 20.52	549.24 ± 20.68	0.351
	ACD	3.16 ± 0.25	3.18 ± 0.24	3.17 ± 0.26	3.17 ± 0.27	3.17 ± 0.26	0.456
	LT	3.56 ± 0.14	3.56 ± 0.14	3.57 ± 0.15	3.56 ± 0.14	3.57 ± 0.14	0.751
	SBP (mmHg)	106.94 ± 10.49	107.76 ± 12.43	107.06 ± 14.99	101.47 ± 11.86	104.06 ± 13.53	**0.048†**
	DBP (mmHg)	72.82 ± 8.68	76.37 ± 12.80	78.82 ± 18.76	72.47 ± 13.90	72.71 ± 13.60	0.153
	SPO2 (%)‡	98(1)	98(1)	98(0)	98(2)	98(0)	0.121
	PI (%)#	3.7(3.2)	1.7(1.9)	2.6(2.4)	1.8(1.9)	1.5(2.2)	**0.005†**

‡ Data deemed as not normal

* p values represent within-group comparisons across time points

† significant p value within-group comparisons across time points

DISCUSSION

This study examined the acute influence of systemic hydration, induced through the water drinking test (WDT), on ocular biometry in myopic and emmetropic eyes. The principal finding was that axial length (AL) in myopic eyes remained stable following water ingestion, whereas emmetropic eyes demonstrated small but statistically significant fluctuations. These results suggest fundamental biomechanical differences between refractive groups, consistent with the adaptive remodeling of scleral tissue in myopia.

The observed AL stability among myopes contrasts with previous reports of transient axial shortening (Read & Collins, 2010) and elongation (Junghans et al., 2013) after WDT, differences that may reflect variation in methodology, sample characteristics, or myopia severity. In the present cohort, the absence of measurable AL change in myopes may reflect altered scleral biomechanics secondary to chronic tissue remodeling. Although progressive myopia is associated with collagen disorganization and reduced intrinsic stiffness (McBrien et al., 2009; Yu & Zhou, 2022), such remodeling also increases viscoelastic damping and reduces the sclera's ability to respond dynamically to short-term pressure fluctuations. Consequently, despite its long-term susceptibility to gradual axial elongation, the myopic sclera may exhibit limited immediate deformation during transient IOP elevations, such as those induced by the water drinking test.

Notably, Yusof et al. (*in press*) confirmed that fluid ingestion elicits a short-term IOP rise in young Malaysian adults. The present study complements that work by demonstrating that, despite the IOP increase expected under similar physiological stress, ocular biometry particularly AL remained unchanged in myopic participants. This implies that the myopic eye, though structurally elongated, behaves as a mechanically stable compartment under transient IOP load, reinforcing the concept of adaptive rather than purely weakened scleral remodeling.

In contrast, the modest AL variability observed in emmetropes may reflect preserved posterior ocular compliance, consistent with the choroid's capacity for rapid volume adjustment in response to hydration and perfusion changes (Zhou et al., 2021). While the fluctuation was not clinically significant, it suggests a more elastic posterior segment capable of dynamic compensation, which likely represents a protective biomechanical feature in non-myopic eyes.

No significant time-dependent effects were observed for CCT, ACD, or LT in either group, indicating that anterior segment structures are minimally influenced by short-term systemic hydration. This finding aligns with the concept of effective ocular autoregulation, wherein corneal and anterior chamber geometries remain stable despite systemic or IOP variations (Luo et al., 2015).

Systemic responses, including increases in systolic blood pressure (SBP) and perfusion index (PI), were consistent with the sympathetic reflex described in previous WDT studies (Jordan et al., 2000; Guidoboni et al., 2014). However, the lack of accompanying biometric change reinforces the eye's resistance to short-term hemodynamic perturbation (Flammer et al., 2013). Collectively, these findings suggest that the posterior segment, particularly in myopia, exhibits stable viscoelastic behavior even when transiently challenged by systemic fluid shifts.

Clinically, this stability is important for interpreting serial axial length data in myopia control. Since hydration-induced variability appears negligible, clinicians can be confident that routine biometry accurately reflects structural progression rather than short-term physiological noise. The consistent AL difference between refractive groups, independent of hydration, underscores the reliability of AL as a biomarker for myopia monitoring.

Future studies should expand sample diversity, incorporate varying myopia severities, and integrate optical coherence tomography (OCT) based choroidal imaging to elucidate the interplay between hydration, choroidal perfusion, and scleral biomechanics. Combining IOP tracking with real-time biometry, building on Yusof et al. (*in press*), may further clarify the dynamic relationship between intraocular pressure fluctuations and structural stability in myopia.

In summary, systemic hydration induced by the WDT produced no significant biometric changes in myopic eyes. Together with evidence from Yusof et al. (*in press*), these findings suggest that although IOP may transiently rise after fluid ingestion, the myopic globe maintains mechanical stability. The inherent rigidity of the remodeled sclera appears to buffer against short-term physiological stress, reaffirming the reliability of AL as a clinical parameter for myopia management.

CONCLUSION

Systemic hydration through the water-drinking test induced short-term systemic changes but did not measurably affect ocular biometry in myopic young adults.

In contrast, emmetropic eyes exhibited small yet statistically detectable axial-length fluctuations, reflecting greater posterior-segment compliance. Overall, these results suggest that the myopic eye maintains mechanical stability under transient physiological stress, consistent with adaptive scleral remodeling rather than structural fragility.

Clinically, this biomechanical stability supports the use of axial length as a reliable indicator of myopia progression, unaffected by short-term hydration or pressure fluctuations. Complementing the transient IOP elevations reported by Yusof et al. (*in press*), the present findings emphasize that functional pressure changes do not necessarily produce structural deformation. Future work integrating real-time biometry with choroidal imaging should further clarify the dynamic interplay among ocular biomechanics, hydration, and intraocular pressure in myopia.

AUTHOR DECLARATIONS

Competing Interests

The authors declare no conflict of interest related to the conduct or publication of this study.

Funding

This study did not receive any specific grant from public, commercial, or not-for-profit funding agencies.

Authors' Contributions

All authors contributed substantially to the conception, design, data collection, analysis, and interpretation of results. A.I.A. and N.S.C.M.Z. conducted data collection, analysis, and initial manuscript drafting. M.A.S.A.R., M.M.S., and N.S.S. verified data analysis, critically revised the manuscript for intellectual content. F.Y. supervised the study, verified data analysis, critically revised the manuscript for intellectual content, and approved the final version for submission. All authors read and approved of the final manuscript.

Acknowledgments

The authors wish to acknowledge the Department of Optometry and Visual Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, for providing research facilities and administrative support.

AI Assistance Disclosure

Portions of this manuscript were refined with the assistance of ChatGPT (OpenAI, GPT-5), used exclusively for language editing and structural improvement under full human intellectual supervision. All conceptualization, data interpretation, and conclusions are the authors' own.

REFERENCES

Flammer, J., Konieczka, K., Bruno, R. M., Virdis, A., Flammer, A. J., & Taddei, S. (2013). The eye and the heart. *European Heart Journal*, 34(17), 1270–1278. https://doi.org/10.1093/eurheartj/eht023

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). *G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Guidoboni, G., Harris, A., Cassani, S., Arciero, J., Siesky, B., Amireskandari, A., Tobe, L., Egan, P., Januleviciene, I., & Park, J. (2014). Intraocular Pressure, Blood Pressure, and Retinal Blood Flow Autoregulation: A Mathematical Model to Clarify Their Relationship and Clinical Relevance. *Investigative Ophthalmology & Visual Science,* 55(7), 4105. https://doi.org/10.1167/iovs.13-13611

Hamdan, N., Sabere, A., Ruslan, A., Buari, N., Abdul Rahim, M., Che Azemin, M., & Yusof, F. (2024). The Short-term Effects of Coffee and Caffeine on Intraocular Pressure in Healthy Subjects. *Journal of Health Science and Medical Research, 42*(6), e20241108. http://dx.doi.org/10.31584/jhsmr.20241108

Holden, B. A., Fricke, T. R., Wilson, D. A., Jong, M., Naidoo, K. S., Sankaridurg, P., Wong, T. Y., Naduvilath, T. J., & Resnikoff, S. (2016). Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. *Ophthalmology,* 123(5), 1036–1042. https://doi.org/10.1016/j.ophtha.2016.01.006

Jordan, J., Shannon, J. R., Black, B. K., Ali, Y., Farley, M., Costa, F., Diedrich, A., Robertson, R. M., Biaggioni, I., & Robertson, D. (2000). The pressor response to water drinking in humans: a sympathetic reflex? *Circulation,* 101(5), 504–509. https://doi.org/10.1161/01.cir.101.5.504

Junghans, B., Chang, C.-J., Siu, A., Soesanto, N., Tu, M., Murphy, M., & Crewther, S. (2013). Intraocular Pressure in Myopic Eyes during the Water Drinking Test. *Investigative Ophthalmology & Visual Science*, 54(15), 1912–1912. https://iovs.arvojournals.org/article.aspx?articleid=2146562

Kim SH, Lee EJ, Han JC, Sohn SW, Rhee T, Kee C. (2016) The Effect of Diurnal Fluctuation in Intraocular Pressure on the Evaluation of Risk Factors of Progression in Normal

Tension Glaucoma. *PLoS One*, 24;11(10):e0164876. https://doi.org/10.1371/journal.pone.0164876.

Luo, X., Shen, Y., Jiang, M., Lou, X., & Shen, Y. (2015). Ocular Blood Flow Autoregulation Mechanisms and Methods. *Journal of Ophthalmology*, 2015, 1–7. https://doi.org/10.1155/2015/864871

McBrien, N. A., Jobling, A. I., & Gentle, A. (2009). Biomechanics of the Sclera in Myopia: Extracellular and Cellular Factors. *Optometry and Vision Science,* 86(1), E23–E30. https://doi.org/10.1097/opx.0b013e3181940669

Morgan, I. G., Rose, K. A., & Ellwein, L. B. (2010). Is emmetropia the natural endpoint for human refractive development? An analysis of population-based data from the refractive error study in children (RESC). *Acta Ophthalmologica,* 88(8), 877–884. https://doi.org/10.1111/j.1755-3768.2009.01800.x

Ohno-Matsui, K., Wu, P.-C., Yamashiro, K., Vutipongsatorn, K., Fang, Y., Cheung, C. M. G., Lai, T. Y. Y., Ikuno, Y., Cohen, S. Y., Gaudric, A., & Jonas, J. B. (2021). IMI Pathologic Myopia. *Investigative Ophthalmology Visual Science*, 62(5), 5. https://doi.org/10.1167/iovs.62.5.5

Raman, P., & Clement, C. (2020). The water drinking: Practical Tips. *Malaysian Society of Ophthalmology Express*, (15), 6-8. https://mso.org.my

Read, S. A., & Collins, M. J. (2010). Water drinking influences eye length and IOP in young healthy subjects. *Experimental Eye Research*, 91(2), 180–185. https://doi.org/10.1016/j.exer.2010.04.015

Yu, Q., & Zhou, J. B. (2022). Scleral remodeling in myopia development. *International Journal of Ophthalmology*, 15(3), 510–514. https://doi.org/10.18240/ijo.2022.03.21

Yusof F. A., Mohammad N. A., Hilmi M. R. (in press). Effects of Fluid Ingestion on Intraocular Pressure in Normal Young Malaysian Adults, *Malaysian J Med Health Sci*, (In press)

Zhou, X., Ye, C., Wang, X., Zhou, W., Reinach, P., & Qu, J. (2021). Choroidal blood perfusion as a potential "rapid predictive index" for myopia development and progression. *Eye and Vision*, 8(1). https://doi.org/10.1186/s40662-020-00224-0

ORIGINAL ARTICLE

Comparative Evaluation of Honey, Aloe Vera, and Olive Oil as Alternative Ultrasound Coupling Agents for Abdominal Imaging

Linati Aina Omar [1], Ummi Farhana Hashim [1,*]

[1]Department of Diagnostic Imaging and Radiotherapy, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
SNR; CNR; Aloe vera; honey; olive oil;

ABSTRACT

Background: Ultrasound imaging requires an effective coupling medium to eliminate air between the transducer and the skin and ensure optimal sound transmission. Commercial ultrasound gel is widely used, but its limited availability in low-resource settings, potential for skin irritation, and risk of contamination highlight the need for accessible natural alternatives. This study aimed to evaluate the image quality produced by aloe vera, honey, and olive oil in comparison with commercial ultrasound gel when imaging the liver, kidney, and aorta using a tissue-mimicking phantom. **Methods:** A comparative experimental design was used to assess four coupling agents: aloe vera, honey, olive oil, and commercial ultrasound gel. All materials were tested under identical imaging conditions using a Kyoto Kagaku ECHOZY Abdominal Set tissue-mimicking phantom and Siemens Acuson X150 system equipped with a 3.5–5 MHz curvilinear transducer. Images of the liver, kidney, and aorta were acquired at fixed ultrasound settings. Quantitative analysis was performed using ImageJ software to measure signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) based on standardised regions of interest. Qualitative image assessment was carried out through a virtual grading analysis (VGA) by three experienced sonographers who independently ranked image sharpness, noise, depth penetration, and overall quality. **Results:** Commercial ultrasound gel produced the highest image quality for the liver, with the highest SNR and CNR values. Aloe vera achieved the best performance for the kidney, producing higher quantitative values compared with the commercial gel. Olive oil demonstrated the highest CNR value for aortic imaging but showed moderate performance in other organs. Honey consistently produced the lowest SNR and CNR values across all organs, indicating limited suitability as a coupling medium. VGA findings aligned with quantitative outcomes, with aloe vera and olive oil performing comparably to commercial gel in selected organs. **Conclusion:** Aloe vera and olive oil demonstrated potential as practical alternatives to commercial ultrasound gel, particularly when standard gel is unavailable. Aloe vera showed the most consistent performance across imaging modalities, while honey was the least effective. These findings support further evaluation of natural products as accessible and cost-effective coupling media, particularly for use in low-resource or mobile ultrasound settings.

INTRODUCTION

Access to commercial ultrasound gel is not always reliable, particularly in rural, low-resource, or mobile healthcare settings where supply shortages, distribution barriers, and cost limitations may restrict its use (Argaw et al., 2022). These constraints can disrupt essential ultrasound services, especially in point-of-care environments. Commercial gels may also contain preservatives such as parabens and phenoxyethanol, which have been linked to skin irritation and allergic reactions in sensitive individuals, and contamination events involving ultrasound gel have been reported (Sherman et al., 2015; Chasset et al., 2015). These challenges underscore the need to identify safe, affordable, and easily obtainable alternatives.

Natural substances such as aloe vera, honey, and olive oil are widely accessible, inexpensive, and commonly used in everyday products. Their favourable biocompatibility (Kaloteraki et al., 2023) and potential acoustic characteristics make them promising candidates for investigation as alternative coupling media, particularly in situations where commercial gel is unsuitable or unavailable).

Although commercial ultrasound gel remains the standard coupling medium (Wray et al., 2023), there is limited empirical evidence evaluating whether natural materials can deliver comparable acoustic performance or diagnostic image quality. Existing studies on natural substances are few, often limited in methodology, and rarely include rigorous quantitative analysis or assessment across multiple abdominal organs under standardised conditions (Paulinus et al., 2022) As a result, there is insufficient evidence to determine whether these alternatives can reliably support diagnostic-quality

* *Corresponding author.*
E-mail address: ummifarahanahashim@iium.edu.my

imaging. A comprehensive evaluation incorporating both quantitative and qualitative image quality measures is therefore needed to address this gap.

The present study sought to identify accessible, cost-effective, and safe options for ultrasound coupling, which was particularly important for low-resource facilities, mobile ultrasound practice, and emergency settings. By evaluating natural materials that were readily available in local markets, the study supported more sustainable and resilient ultrasound services. If certain natural substances were shown to provide image quality comparable to commercial gel, they could have enhanced continuity of care, reduced dependency on manufactured consumables, and expanded the availability of diagnostic imaging in low-resource areas.

This study aimed to evaluate the potential of aloe vera, honey, and olive oil as alternative ultrasound coupling agents by thoroughly comparing them with commercial ultrasound gel. The investigation included a detailed quantitative analysis using SNR and CNR measurements to assess transmission efficiency, as well as a qualitative evaluation through VGA by experienced sonographers.

MATERIALS AND METHODS

This study was conducted at the Diagnostic Imaging Laboratory, Department of Diagnostic Imaging and Radiotherapy, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Bandar Indera Mahkota, Kuantan, Pahang.

This comparative experimental study was designed to assess the performance of three natural substances which are aloe vera, honey, and olive oil as potential alternatives to commercial ultrasound gel. The commercial gel was used as the reference standard, and all coupling media were evaluated under identical imaging conditions using an abdominal ultrasound phantom. Image quality was assessed through SNR, CNR and VGA to determine their diagnostic suitability.

Coupling Media

Four ultrasound coupling media were selected for this study based on their availability, purity, and suitability for direct application without further preparation. These included a commercial ultrasound gel (Sky Gel Ultrasound Transmission Medical Gel); an aloe vera gel (Fruit of the Earth 100% Aloe Vera Gel); a high-purity natural honey (SIHAT Sidr Yemeni Honey); and a cold-pressed extra virgin olive oil (SIHAT, Syria). These materials were chosen because they are readily obtainable in retail outlets, making them practical options for use in resource-limited or emergency ultrasound settings.

Experimental Setup

Ultrasound imaging was performed using a Siemens Acuson X150 system equipped with a 3.5–5 MHz curvilinear transducer, which is appropriate for abdominal organ evaluation. A Kyoto Kagaku ECHOZY Abdominal Set tissue-mimicking phantom was used to replicate abdominal structures, including the liver, right kidney, and aorta.

Before imaging, the transducer was prepared to ensure consistent acoustic transmission. A thin layer of commercial gel was applied to the probe surface, after which it was tightly wrapped in plastic food wrap to prevent direct contact with the test substances while maintaining adequate echo transmission. The phantom's surface was also wrapped with plastic film to avoid contamination. The transducer was cleaned and rewrapped after every coupling media to maintain experimental integrity.

Ultrasound Scanning Parameters

A frequency of 3.6 MHz was used for all three organs as it provides an optimal balance between image detail and depth penetration, making it suitable for abdominal imaging and ensuring adequate visualisation of the liver, kidney, and aorta.

Gain settings were adjusted according to organ characteristics to optimise image brightness. The liver was imaged at 55 dB to enhance visibility of its relatively homogeneous and moderately echogenic parenchyma. In kidney, the gain was reduced to 52 dB to prevent oversaturation of the mixed echogenic renal cortex and medulla. The aorta required the lowest gain, 45 dB, to clearly differentiate the vessel wall and lumen without over-amplifying echoes.

Dynamic range (DR) values were also tailored for each organ. The liver was imaged with a DR of 55 dB to preserve subtle parenchymal differences. A narrower DR of 50 dB was used for the kidney to increase contrast between the cortex, medulla, and collecting system. In contrast, the aorta was assigned a higher DR of 60 dB to capture the full range of signals from the hyperechoic wall to the anechoic lumen.

A persistence value of 3 was applied uniformly across all organs to reduce speckle noise while preserving fine structural details, and an edge-enhancement setting of 1 was used to provide mild sharpening of organ boundaries.

The resolution/scale parameter was standardised at 3 to maintain consistent image quality across organs. All images were displayed using a grayscale map to ensure uniform contrast for accurate SNR and CNR calculations. A tint value of 1 was applied to prevent any colour bias and maintain a neutral grayscale appearance, supporting reliable quantitative and qualitative image assessment.

Scanning Technique

The Kyoto Kagaku ECHOZY Abdominal Set tissue-mimicking phantom was scanned several times for each organ and each coupling medium to ensure consistency in image acquisition, with the transducer maintained at the same location and imaging plane for every scan. The most representative image, defined as the clearest and most stable image obtained under fixed parameters, was then selected for both quantitative SNR and CNR analysis and qualitative VGA assessment.

Quantitative Analysis: Signal-to-Noise Ratio and Contrast-to-Noise Ratio

Quantitative assessment of image quality was performed using ImageJ software to calculate the SNR and CNR for all images obtained with each coupling medium. SNR and CNR provide measurements of image clarity, tissue contrast, and background noise which are the key determinants of diagnostic usefulness in ultrasound imaging.

All ultrasound images were exported in high-resolution format and analysed using a standardised region-of-interest (ROI) approach. Consistency in ROI placement was maintained across all images to ensure accurate comparisons. ROI placement was standardised by using fixed anatomical reference points within the phantom, and ImageJ software enabled the same ROI shapes and sizes to be replicated across all images. The evaluator applied identical ROI templates to corresponding anatomical regions for every coupling medium, ensuring that all SNR and CNR measurements were derived from comparable and reproducible locations.

Descriptive Analysis of SNR and CNR

Data were presented as mean values with their corresponding standard deviations. Due to the limited number of repeated scans, inferential statistical testing was not conducted. As a result, comparisons across coupling media are descriptive in nature and should be interpreted carefully.

Liver region-of-interest placement

In liver, three rectangular ROIs were placed within the liver parenchyma to assess tissue uniformity, while three circular ROIs were positioned along the diaphragm (Figure 1). The diaphragm serves as a crucial reference point due to its strong echogenic contrast against the liver tissue, allowing for a clearer evaluation of boundary definition and sound wave penetration. These ROI placements enable a comprehensive assessment of liver echogenicity by capturing variations between soft tissue structures and high-contrast interfaces (Theodotou et al., 2019).

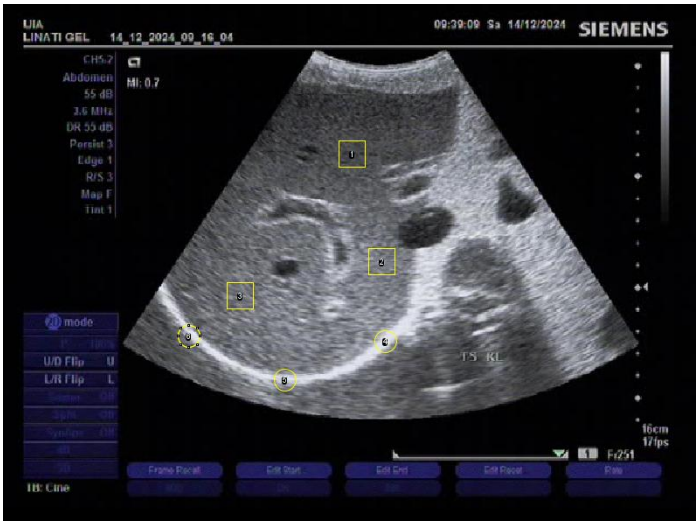


Figure 1: Ultrasound Image of the Liver with ROI Markings

Kidney region-of-interest placement

In right kidney, three ROIs were placed over the renal cortex, medulla, and sinus to represent its key anatomical structures (Figure 2). The cortex and medulla help evaluate tissue contrast and homogeneity, while the sinus, which contains fat and vascular elements, reflects the acoustic penetration and signal (Beutler et al., 2024).

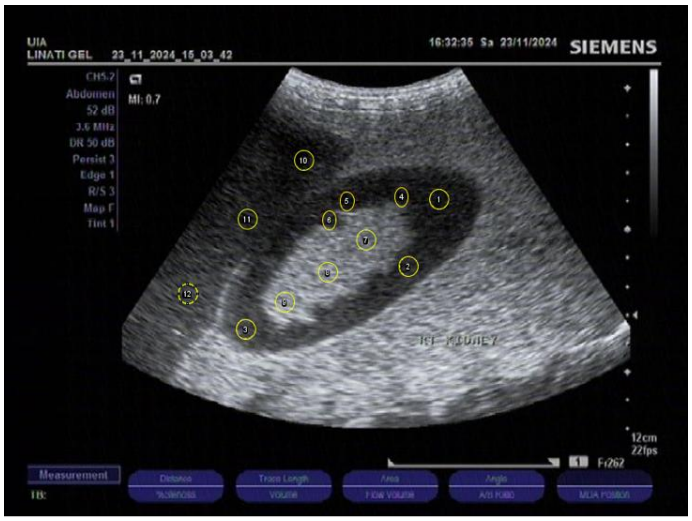


Figure 2: Ultrasound Image of the Kidney with ROI Markings

Aorta region-of-interest placement

In aorta, a rectangular region of interest was placed within

the lumen, while the mean wall intensity was derived from several point measurements taken along the delineated vessel wall (Figure 3). Three narrow rectangular background ROIs were positioned adjacent to the aorta to represent surrounding vascular tissue and to calculate the standard deviation (SD) of the background noise.

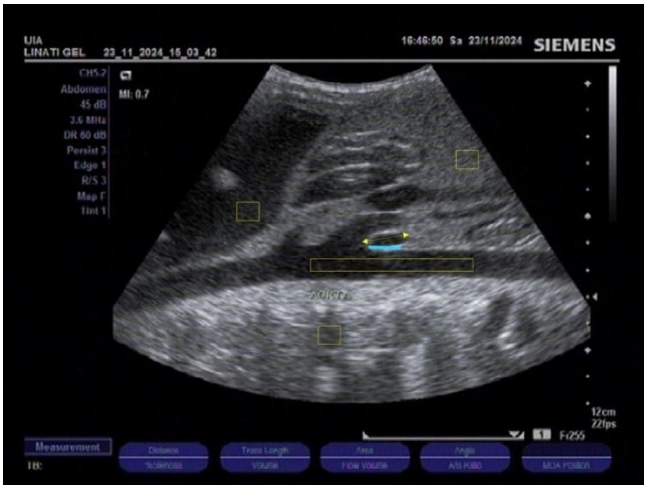


Figure 3: Ultrasound Image of the aorta with ROI Markings

Signal-to-Noise Ratio Calculation

The SNR was calculated to determine the relationship between the signal intensity of the target tissue and the degree of noise in the surrounding background. Higher SNR values indicate clearer images with reduced noise interference. SNR was computed for each ROI using the formula:

$$\text{SNR} = \frac{\mu_{\text{ROI}}}{\sigma_{\text{BG}}} \quad (1)$$

where, μ_{ROI} = mean pixel intensity of the tissue of interest
σ_{BG} = SD of pixel intensities within the background ROIs

Contrast-to-Noise Ratio Calculation

The CNR was calculated to evaluate how effectively each coupling medium allowed differentiation between adjacent anatomical structures. The general formula used was:

$$\text{CNR} = \frac{\mu_{\text{ROI}} - \mu_{\text{BG}}}{\sigma_{\text{BG}}} \quad (2)$$

where, μ_{ROI} = mean pixel intensity of the tissue of interest
μ_{BG} = mean pixel intensity of the background
σ_{BG} = SD of pixel intensities within the background ROIs

Qualitative analysis: Virtual Grading Analysis

Qualitative image quality was assessed using a VGA approach, with the evaluation criteria adapted from Smit and Breedt (2022). Three sonographers with more than ten years of diagnostic ultrasound experience independently reviewed the images.

Images of the liver, right kidney, and aorta obtained using using different coupling medium were exported in digital format and organised into organ-specific sets. Each set contained four images, one representing each coupling medium. To minimise observer bias, the images were anonymised and relabelled as Image A, B, C, and D, without disclosing the coupling agent used.

The VGA utilised a structured scoring form comprising four criteria: (1) detail and sharpness, (2) image noise, (3) depth and penetration, and (4) overall image quality. The sonographers ranked the four images on a scale from 1 to 4 for every organ and criterion, where 1 indicated the highest quality and 4 the lowest. This ranking process resulted in 12 rank values for each image (4 criteria × 3 reviewers), derived from four criteria assessed by three reviewers using a forced-ranking format. The mean rank was calculated by summing all 12 assigned ranks and dividing the total by 12, following the formula:

$$\text{Mean Rank} = \frac{Sum\ of\ all\ assigned\ ranks\ for\ that\ image}{12} \quad (3)$$

RESULTS

Quantitative Analysis: Signal-to-Noise Ratio and Contrast-to-Noise Ratio

Liver imaging results (Table 1) show that commercial ultrasound gel produced the best image quality, with the highest SNR (2.78) and CNR (1.93), indicating superior signal clarity and structure definition. Olive oil performed closely behind (SNR 2.72; CNR 1.68), suggesting it could serve as a practical alternative when commercial gel is not available. Aloe vera showed slightly lower image clarity (SNR 2.63), and honey produced the poorest results (SNR 1.36; CNR 1.55), reflecting weak contrast and limited visualisation of liver structures.

In kidney imaging, aloe vera outperformed all other coupling agents, with the highest CNR (3.37) and SNR (5.82). These values exceeded those of commercial gel (CNR 2.27; SNR 4.71), indicating that aloe vera may be a strong alternative for renal assessment. Olive oil produced moderate-quality images (SNR 3.19; CNR 1.37), while honey again delivered the lowest values (SNR 2.19; CNR 0.61), showing poor structural detail and signal quality.

Aortic imaging showed a different trend, with olive oil producing the highest CNR (6.82), providing the best contrast between the aortic wall and lumen. However, its SNR (1.21) was slightly lower than the commercial gel (1.78), meaning it had weaker overall signal clarity. Aloe vera yielded moderate performance (SNR 1.12; CNR 3.76), and honey again produced the lowest values (SNR 0.21; CNR 2.86), indicating visibly poor contrast and unsuitable image quality for aortic evaluation.

Table 1: SNR and CNR for different coupling medium and organ

Part	Coupling Medium	SNR	CNR
Liver	Aloe Vera	2.63	1.59
	Honey	1.36	1.55
	Olive Oil	2.72	1.68
	Ultrasoud Gel	2.78	1.93
Kidney	Aloe Vera	5.82	3.37
	Honey	2.19	0.61
	Olive Oil	3.19	1.37
	Ultrasoud Gel	4.71	2.27
Aorta	Aloe Vera	1.12	3.76
	Honey	0.21	2.86
	Olive Oil	1.21	6.82
	Ultrasoud Gel	1.78	5.12

Qualitative analysis: Virtual Grading Analysis

Based on Table 2, honey produced the best perceived image quality for the liver, with average rank of 2.17. This means the sonographers consistently judged honey to provide sharper images with less noise and good penetration. Commercial ultrasound gel followed closely with a score of 2.25, indicating strong and clinically acceptable performance. Aloe vera received an average rank of 2.58, suggesting moderate but acceptable image quality, while olive oil ranked lowest at 3.00, indicating reduced sharpness, more noise, and poorer depth penetration for liver imaging.

While for the kidney, aloe vera achieved the highest-quality images with an average rank of 1.83 as it provides the clearest visibility of renal structures, including the cortex, medulla, and sinus. Commercial ultrasound gel was the second-best performer with an average rank of 2.00, producing high-quality images. Olive oil showed moderate performance with a ranking of 2.83, while honey performed the worst with a rank of 3.33, indicating noticeably higher noise and reduced contrast, making it the least effective medium for kidney imaging.

Evaluation of the aorta showed that aloe vera again produced the best perceptual image quality, ranking first with an average score of 1.67. Aloe vera provided clear vessel wall definition, good lumen visibility, and overall balanced contrast. Honey ranked second with an average score of 2.17, indicating that while its performance was acceptable. Olive oil received a score of 3.00, reflecting moderate clarity with some limitations in contrast and noise. Commercial ultrasound gel ranked last with 3.17, suggesting that the sonographers judged its aortic images to be the least sharp and least distinct compared with the natural alternatives.

Table 2: VGA for different coupling medium and organ

Part	Average Rank			
	Aloe Vera	Honey	Olive Oil	Ultrasoud Gel
Liver	2.58	2.17	3.00	2.25
Kidney	1.83	3.33	2.83	2.00
Aorta	1.67	2.17	3.00	3.17

DISCUSSION

The quantitative analysis indicates that although commercial ultrasound gel performed strongly overall, it did not consistently produce the highest SNR and CNR values across all organs assessed. While it remains the clinical standard and delivered excellent image quality for liver imaging, its performance was exceeded by aloe vera in kidney evaluation and by olive oil in aortic contrast assessment. These outcomes suggest that commercial gel continues to provide reliable acoustic transmission, as reported in previous studies (Williams et al., 2019), but its advantage is not uniform across all anatomical regions.

Among the alternative materials tested, olive oil showed notable potential, particularly in vascular imaging. Its exceptionally high CNR for the aorta indicates enhanced differentiation between the vessel wall and lumen, which is crucial for accurate assessment of vascular structures. Olive oil also demonstrated competitive performance for liver imaging, ranking close to the commercial gel. The improved contrast observed in this study suggests that olive oil may offer diagnostic advantages in settings where commercial gel is unavailable.

Aloe vera showed its strongest performance in kidney imaging, surpassing the commercial gel in both SNR and CNR. This finding indicates that aloe vera's acoustic behaviour may align particularly well with the tissue composition of the kidney, enabling clearer structural differentiation. The superior performance in renal imaging suggests that aloe vera may be a viable natural alternative for specific organ applications.

In contrast, honey delivered consistently poor results for all organs, showing the lowest SNR and CNR values across the dataset. Its high viscosity likely impairs efficient sound transmission, resulting in reduced signal quality and weak structural contrast. The consistently suboptimal performance suggests that honey is unsuitable as a general-purpose ultrasound coupling agent.

Qualitative assessment using VGA provided additional insight into the perceptual quality of images produced by each coupling medium. Aloe vera ranked highly for both the right kidney and aorta, reflecting strong performance across multiple evaluators. Its favourable viscosity and acoustic behaviour appear to reduce image noise while improving boundary clarity, leading to consistently high-quality visualisation. The highlight its potential as a practical substitute when commercial gel is not available.

Honey exhibited mixed qualitative performance. Although it scored well for the liver, its ranking for the kidney was substantially poorer. This pattern suggests that honey may interact differently depending on organ type or tissue echogenicity. While certain physical properties might have temporarily enhanced liver image appearance, its consistently low SNR and CNR across the quantitative analysis indicate that honey does not reliably support diagnostic-quality imaging. These inconsistent outcomes align with previous findings that highlight the need to consider anatomical variation when evaluating unconventional coupling agents (Okere, 2019).

Olive oil showed moderate performance across most organs, receiving intermediate rankings that reflect acceptable but not superior image quality. While its strong quantitative contrast performance for the aorta was promising, the qualitative results suggest that olive oil may not provide uniform clarity across all anatomical regions.

Commercial ultrasound gel, despite being the reference standard, showed varied rankings in the VGA. Its strong performance for the right kidney contrasted with its lower rank for aortic imaging. This inconsistency suggests that while commercial gel is generally reliable, perceptual image quality may still vary according to specific anatomical or operator-related factors.

CONCLUSION

This study demonstrates that while commercial ultrasound gel remains a reliable reference standard, its performance was not consistently superior across all organs. Quantitative assessment showed that commercial gel produced the best liver image quality but was outperformed by aloe vera in kidney imaging and by olive oil in aortic contrast evaluation. These findings indicate that certain natural materials may offer organ-specific advantages in acoustic transmission.

Among the alternatives, olive oil showed strong potential, producing the highest aortic CNR and competitive liver imaging quality, suggesting suitability for vascular and soft-tissue applications when commercial gel is unavailable. Aloe vera delivered the clearest kidney images in SNR, CNR and ranked highly in qualitative assessments, demonstrating favourable acoustic behaviour and consistent visual clarity across evaluators.

Honey performed poorly across most measurements, with the lowest SNR and CNR values and weak qualitative rankings, indicating limited suitability as a coupling medium.

Taken together, the findings suggest that aloe vera and olive oil may be considered as provisional alternatives in specific clinical scenarios, particularly in resource-limited settings. Their reliable performance highlights the potential value of natural coupling agents in expanding ultrasound accessibility when commercial gel is restricted or unavailable. However, as this was a phantom-based study, the results may not fully replicate in clinical environments, and further in vivo studies are required before these natural media can be recommended for routine clinical use.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Argaw, M., Abawollo, H., Tsegaye, Z., Beshir, I., Damte, H., Mengesha, B., & Desta, B. (2022). Experiences of midwives on vscan limited obstetric ultrasound use: a qualitative exploratory study. *BMC Pregnancy and Childbirth, 22(1).* https://doi.org/10.1186/s12884-022-04523-3

Beutler, B. D., El-Sabawi, B., Walker, D. K., Cen, S., & Tchelepi, H. (2024). An Objective Computer-Assisted Measurement of Sonographic Renal Cortical Echogenicity: The Splenorenal Index. *Ultrasound quarterly, 40(1), 56–60.* https://doi.org/10.1097/RUQ.0000000000000646

Chasset, F., Soria, A., Moguelet, P., Mathian, A., Auger, Y., Françès, C., & Barète, S. (2015). Contact dermatitis due to ultrasound gel: a case report and published work review. *The Journal of Dermatology, 43(3), 318-320.*

https://doi.org/10.1111/1346-8138.13066

Kaloteraki, C., Bousdouni, P., Almpounioti, K., Ouzaid, C., Papagianni, O., Sfikti, F., & Koutelidakis, A. E. (2023). Fortification of olive oil with herbs and waste by-products towards sustainable development: total antioxidant capacity, phenolic content, and in vitro predicted bioavailability. *Applied Sciences, 13(15), 8876*. https://doi.org/10.3390/app13158876

Okere, P. C. (2019). Low-cost antimicrobial fortification of ultrasound coupling gel: an ergonomic innovation to combat sonology-acquired nosocomial infections. *Malawi Medical Journal, 31(1), 45.* https://doi.org/10.4314/mmj.v31i1.8

Paulinus, S. O., Egong, A. E., Ukpong, E. V., Erim, A. E., Eru, E. M., Ojile, J. A., Efanga, I., & Egbe, N. O. (2022b). Aloe vera Lotion Used as a Potential Alternative Couplant for Sonography, in a Low-Resource Health Care Facility. *Journal of Diagnostic Medical Sonography, 39(2), 116–121.* https://doi.org/10.1177/87564793221120430

Sherman, T., Ferguson, J., Davis, W., Russo, M., & Argintar, E. (2015). Does the use of ultrasound affect contamination of musculoskeletal injections sites?. *Clinical Orthopaedics and Related Research, 473(1), 351-357*. https://doi.org/10.1007/s11999-014-3903-4

Smit, C., & Breedt, A. (2022). Petroleum jelly as an alternative coupling medium in focus assessed transthoracic echocardiography (FATE). *Journal of Anesthesia and Intensive Care Medicine. 2022;11(5).* https://doi.org/10.19080/JAICM.2022.11.555825.

Theodotou, M., Fokianos, K., Moniatis, D., Kadlenic, R., Chrysikou, A., Aristotelous, A., Mouzouridou, A., Diakides, J., & Stavrou, E. (2019). Effect of resveratrol on non-alcoholic fatty liver disease. *Experimental and therapeutic medicine, 18(1), 559–565.* https://doi.org/10.3892/etm.2019.7607

Williams, G., Cai, C., Artime, C., Teima, D., Emerald, A., & Rajkumar, K. (2019). Evaluation of saline versus gel as contact medium for ultrasonographic image quality. *Journal of Clinical Ultrasound, 47(9), 540-545.* https://doi.org/10.1002/jcu.22750

Wray, J., Soucy, Z., Daniel, N., Weinberg, N., Krauthamer, G., Crockett, S., & Storn, J. (2023). Comparison of commonly carried liquids against commercial ultrasound gel for use in the backcountry setting. *Wilderness and Environmental Medicine, 34(2), 135-142.* https://doi.org/10.1016/j.wem.2022.12.005

ORIGINAL ARTICLE

Mental Health Status, Influencing Factors and Self-Coping Approaches Among Medical Imaging Students at International Islamic University Malaysia

Muhammad Putera Hakimi Mohd Ghazali[1], Norhanna Sohaimi[1,*], Fatihah Syafinaz Kamarul Zaman[1], Umi Nadrah Amran[1]

[1]Department of Diagnostic Imaging and Radiotherapy, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia

Keywords:
mental health; influencing factor; self-coping; medical imaging; student

ABSTRACT

Background: This research investigates the mental health status, influencing factors, and self-coping strategies among Medical Imaging students at the International Islamic University Malaysia (IIUM). Recognizing the increasing trend of mental health issues among students, this study aims to identify the prevalence of depression, anxiety, and stress, and to explore the factors contributing to these conditions. **Methods:** The study used a cross-sectional design. The target population was IIUM Medical Imaging students (Year 1 to Year 4). Convenience sampling was performed. The required sample size was 79, calculated using the Krejcie and Morgan method. A total of 88 completed questionnaires were returned. Data collection was conducted online from March to May 2024, using a newly designed questionnaire consisting of four parts, including the Depression, Anxiety, and Stress scale (DASS-21). Data analysis for the objectives used descriptive analysis. **Results:** A cross-sectional study (n=88) among Medical Imaging students, predominantly female (76.1%), revealed significant mental distress, with anxiety being the most prevalent concern, as only 37.5% (n=33) of students reported normal anxiety levels, and combined severe and extremely severe anxiety affected 32.9% of the cohort. Stressors exhibited marked gender differences: male students primarily cited financial and family issues (76.2% for both) as major factors, whereas female students most frequently reported study-related stress (73.1%), self-related issues (73.1%), and peer relationships (71.6%). Regarding coping, both genders favoured passive and spiritual approaches, such as sleeping (Male: 80.9%; Female: 80.6%), but males reported higher engagement in physical activity (57.1% vs. 20.9% for females), while females preferred social/creative outlets. **Conclusion:** The study underscores the critical need for increased awareness and interventions to address mental health issues among university students. It highlights the importance of providing adequate support systems and promoting effective self-coping mechanisms to improve the overall well-being and academic performance of students. The research contributes valuable insights into the mental health landscape of Medical Imaging students and offers recommendations for future studies and institutional policies to enhance student mental health support.

INTRODUCTION

The Global and National Burden of Mental Health Disorders

Mental health is fundamental to human well-being, influencing an individual's cognitive abilities, performance, and social interactions. Currently, mental health disorders represent a substantial global public health burden. According to the World Health Organization (WHO), approximately one in seven individuals globally, equivalent to about 1.1 billion people was affected by a mental disorder in 2021, with anxiety and depressive disorders identified as the most prevalent conditions (Institute for Health Metrics and Evaluation, 2024). This issue is acutely reflected in the Malaysian context, particularly among the youth population. Recent evidence from the Institute for Youth Research Malaysia (IYRES) and the United Nations Children's Fund (UNICEF) indicates a substantial mental health burden among Malaysian youth. Findings from the Malaysian Youth Mental Health Index 2023 (MyMHI'23) reported that approximately six in ten individuals aged 15 to 40 exhibited mild to severe depressive symptoms (IYRES & UNICEF, 2024). The same assessment documented an overall mental health score of 71.91, reflecting a moderate risk level across the youth population (The Star, 2024). Complementary analyses further revealed that nearly one in ten Malaysian youths had experienced suicidal ideation, plans, or attempts as early as 2020 (IYRES & UNICEF, 2024).

The situation has been further aggravated by the COVID-19 pandemic. National surveillance data reported a nearly fivefold rise in mental-health-related help-seeking in 2022 compared to 2020, suggesting a worsening trend in psychological distress (Sulaiman, 2022). Collectively, these data highlight an escalating mental health crisis among Malaysian youth and underscore the need for focused,

* *Corresponding author.*
E-mail address: norhanna@iium.edu.my

evidence-based research and targeted interventions for vulnerable population groups.

Academic Stress, Burnout, and the Impacts

The rising prevalence of mental health issues among young adults (Jurewicz, 2015) is strongly correlated with stressors originating from their educational environment, including financial difficulties, learning challenges, and particularly, academic workload. The rigorous academic demands and high-stakes evaluation inherent in university education are key contributors to student burnout, a syndrome linked to negative educational attitudes and recognized within academic contexts (Liu et al., 2023).

Academic workload and separation from school were found to have negative effects on college students' health via perceived stress (Yang et al., 2021). Inadequate sleep is also associated with poor mental health outcomes (Milojevich & Lukowski, 2016; Yaghmour et al., 2023). Additionally, diet routine and obesity can influence mental illness by impacting mood, and improvements in diet quality have been linked to better mental health outcomes (Jacka et al., 2011).

This intense academic stress is a major concern, as it has been empirically linked to the development of serious mental health problems such as depression, anxiety, and substance abuse (Pascoe et al., 2020, Kadhum et al., 2022). If left unaddressed, academic stress and subsequent burnout can profoundly impair a student's emotional regulation, psychological resilience, and overall capacity to succeed in their academic and personal life.

Poor mental health has far-reaching consequences across multiple domains. It can lead to chronic physical conditions like cardiovascular disease, diabetes, and obesity, emphasizing the intricate, bidirectional relationship between mental and physical health. In social relationships, conditions like depression and anxiety frequently lead to social isolation and loneliness, impeding recovery (Hamza et al., 2021).

Student-Specific Challenges, Coping, and Assessments

Students face unique challenges, with medical students, for example, showing significantly higher rates of depression compared to the general population due to their workload (Slavin et al., 2014). University life exposes students to challenges that can improve their resilience and adaptation to the adult world. Self-coping approaches are critical for managing stress, utilizing methods like mindfulness, physical exercise, and in the Islamic perspective, performing prayer and reciting the Quran.

For assessment, the Depression Anxiety Stress Scales-21 (DASS-21) measures emotional states (Laranjeira et al., 2023), while the Perceived Stress Scale-10 (PSS-10) assesses the perception of stress (Sandu et al., 2015). The Depression Anxiety Stress Scales-21 (DASS-21) and the Perceived Stress Scale-10 (PSS-10) are psychometric instruments that measure distinct facets of psychological distress, offering complementary data for comprehensive mental health assessment. The DASS-21 functions as a state-based measure, providing a quantitative score across three primary emotional states: depression, anxiety, and stress, by assessing symptomatic manifestations over a recent period. Conversely, the PSS-10 is designed as a measure of perceived stress, focusing on the cognitive appraisal of life circumstances; specifically, it assesses the degree to which individuals judge their lives to be unpredictable, uncontrollable, and overwhelming. Therefore, the DASS-21 quantifies the severity of internal emotional symptoms, while the PSS-10 quantifies an individual's psychological judgment of their external stress exposure.

Study Rationale and the Focus on IIUM Medical Imaging Students

Given the documented vulnerability of university students and the increasing trend of mental health challenges in Malaysia, this study targets IIUM Medical Imaging undergraduate students. This cohort represents young adults navigating the academic pressures of a demanding professional field, making them susceptible to stress factors such as peer pressure, complex subject matter, heavy assessment schedules, and time constraints. Crucially, the study also seeks to investigate their self-coping approaches.

As a population within an Islamic institution, their coping repertoire is expected to include both conventional and faith-based strategies, such as worship to Allah SWT, seeking advice from family and friends, alongside activities like social media use, entertainment, sports, and art. Understanding these coping mechanisms is vital, as effective strategies are key determinants of positive mental health outcomes. The aim of this study is, therefore, to systematically identify the mental health status, contributing influence factors, and the self-coping approaches utilized by these students.

The general objective of this study was to investigate the mental health status, influencing factors, and self-coping approaches among IIUM Medical Imaging students. This overarching goal was supported by three specific objectives: (i) To identify the mental health status of the

students using the Depression, Anxiety and Stress Scale 21 (DASS-21); (ii) To identify the influencing factors relating to mental health among the students; and (iii) To explore life satisfaction and self-coping approaches among the students, with particular focus on gender differences.

MATERIALS AND METHODS

The study employed a cross-sectional research design, which involves collecting data from a population or subset at a single time point to provide a snapshot of student's mental health status, influencing factors and self-coping approaches.

Population and Sampling Design

The target population for this research was Medical Imaging students at the International Islamic University Malaysia (IIUM), including all male and female students from Year 1 until Year 4.

Inclusion and Exclusion Criteria

The selection of participants was based on the following criteria as tabulated in Table 1.

Table 1: Inclusion and Exclusion Criteria for Participant Selection

Category	Inclusion	Exclusion
Participants	IIUM Medical Imaging students	Other than IIUM Medical Imaging students
Study Status	Status of study should be active	Not active in study status
Head Injury	No history of severe head injury	History of severe head injury
Medication	Do not take any medication related to mental health	Currently taking any medication related to mental health
Medical History	No history of mental health, psychiatry, etc.	History of significant medical illness or psychiatric/ neurological disorder

Sample Size and Sampling Method

A total of 99 IIUM Medical Imaging students were approached as participants. The required sample size was determined to be 79 students using the Krejcie and Morgan method (Krejcie & Morgan, 1970) and Equation 1. The equation is defined as:

$$s = X^2\, NP(1-P)/[d^2\,(N-1) + X^2\, P(1-P)] \quad (1)$$

where, s = The required sample size
X^2 = The table value of chi-square for 1 degree of freedom at the desired confidence level (3.841)
N = The population size (99)
P = The population proportion (assumed to be 0.50)
D = The degree of accuracy expressed as a proportion (0.05).

Out of the 99 students who were invited to participate, 88 students returned completed questionnaires, yielding a response rate of 88.9%. The sampling method utilized was convenience sampling, also known as availability sampling. This is a non-probability sampling technique where participants are chosen based on their convenient availability and accessibility to the researcher.

Study Location and Duration

The study was conducted at the International Islamic University Malaysia (IIUM), Kuantan Campus, involving students from the Department of Diagnostic Imaging and Radiotherapy (Year 1 to Year 4). The overall study, including the research proposal, ethical clearance, data collection, data analysis, and thesis write-up, was conducted from October 2023. The data collection specifically took place from March 2024 until May 2024.

Ethical Approval

Two levels of ethical clearance were obtained, following the tenets of the Declaration of Helsinki; i) Kulliyyah Postgraduate and Research Committee (KPGRC), Kulliyyah of Allied Health Sciences (Approved 24th January 2024) and, ii) IIUM Research Ethical Committee (IREC) (Approved 20th February 2024). These approvals were compulsory prior to commencing the research, as the study involved human participants, even though no physical intervention was needed.

Questionnaire Design and Data Collection

A set of questionnaires was designed to assess the prevalence of depression, anxiety, and stress related to coping styles and social supports. The questionnaire was divided into four parts: Part A, B, C and D.

Part A consist of socio-demographic items which assessed variables like email address, gender, year of study, age, health status, and history of diseases and medication. Anonymity was successfully maintained using a data

separation protocol, despite the necessary collection of identifying information (e.g., email address) for crucial procedural purposes. The identifiers were collected strictly to prevent submission redundancy and to verify participant eligibility against the study's inclusion criteria. Following these verification checks, the identifiers were stripped from the main analysis file, ensuring the final, aggregated dataset contains only de-identified responses and preserving the integrity of participant privacy. Part B was the validated Depression, Anxiety, and Stress scale (DASS-21) questionnaire. The DASS-21 uses a 4-point Likert scale (0–3) where respondents rate how much each statement applied to them over the past week, with 0 = did not apply at all, 1 = applied some of the time, 2 = applied a good part of the time, and 3 = applied most of the time. Each of the three subscales—Depression, Anxiety, and Stress—contains seven items, and the scores for each subscale are summed and then multiplied by 2 to align with the original 42-item version. Higher scores indicate greater severity of symptoms.

While Part C comprised questions on possible influencing factors questionnaire which covered factors potentially affecting student mental health, such as family, financial issues, peers, life events, study-related problems, and trauma incidents. And Part D focused on self-coping approaches to assess various self-coping approaches like exercising, listening to music, sleeping, and worshiping Allah SWT and life satisfaction questionnaire; Satisfaction with Life Scale (SWLS) which is a 5-item instrument utilizing a 7-point Likert scale (ranging from 1= Strongly Disagree to 7= Strongly Agree) to assess subjective well-being (Townshend, 2025). This scale is designed to quantify an individual's global cognitive judgment regarding their overall life satisfaction, functioning as a reliable measure of this construct as a single, unidimensional factor.

The data was collected using an online-based questionnaire distributed via an online platform, Google Form. Data collection was carried out during the active semester, starting from Week 2 of IIUM Academic Calendar. Participants were made anonymous to respect privacy, and the data was solely for research purposes.

Data Analysis

Descriptive analysis was used to analyse all the specific objectives of the study. These objectives included studying the mental health status, identifying possible influence factors relating to mental health, and exploring life satisfaction and self-coping approaches among medical imaging students based on gender.

RESULTS

Demographic Data

The total number of completed questionnaires was 88. The demographic analysis of the Medical Imaging students demonstrated a significant gender imbalance in the cohort, comprising 67 females (76.1%) and 21 males (23.9%). The distribution of respondents across the four years of study was approximately even, with Year 4 having the highest proportion (27.3%, n=24) and Year 1 and Year 3 each accounting for 23.9% (n=21 each). In terms of health, most students were not on medication (90.9%, n=80) or afflicted with any disease (87.5%, n=77). The few reported medical conditions and medications were confirmed not to be related to mental health issues.

Depression, Anxiety, and Stress Scale (DASS-21)

Prevalence by Year of Study

Analysis of the DASS-21 scores by academic year (Table 2) showed that most students fell within the normal range for stress (61.4%, n=54) and depression (52.3%, n=46). However, anxiety presented a more concerning trend, with only 37.5% (n=33) of students rated as having normal anxiety levels, the lowest percentage across all three categories.

In terms of severe mental distress, 9.0% (n=8) of students reported severe depression, with the highest prevalence in Year 3. Extremely severe anxiety affected 17.0% (n=15) of students, with Year 4 showing the highest count (n=4). Severe anxiety affected 15.9% (n=14) of students, with the highest prevalence in Year 1 (n=7).

Prevalence by Gender

When examining DASS-21 scores by gender (Table 2), a greater number of females reported mild stress (n=12) compared to males (n=4). Conversely, all three cases of extremely severe depression were reported by male students, with no females reporting this severity level. For anxiety, the distribution of severe and extremely severe cases was approximately equal, affecting 17.0% (n=15) of the cohort for each category, indicating significant anxiety across both genders.

Overall, gender-based DASS-21 data indicate that, while both male and female students experience stress, depression, and anxiety, the severity and distribution of these conditions differ. Females are more likely to report mild to moderate stress and depression, whereas males report more cases of severe depression. Anxiety affects both genders, with severe and extremely severe cases being equally prevalent. These findings highlight the importance of gender-specific mental health support in

addressing the unique challenges that male and female medical imaging students face.

Table 2: The DASS-21 data based on the year of study and gender of the Medical Imaging students

	Year of Study, *n*				Gender, *n*	
	1	**2**	**3**	**4**	**Male**	**Female**
Stress						
Normal	14	13	10	17	11	43
Mild	3	4	6	4	4	12
Moderate	4	2	3	2	6	5
Severe	-	3	1	1	-	5
Extremely severe	-	-	1	-	-	1
Depression						
Normal	12	12	8	14	11	36
Mild	2	1	1	3	2	5
Moderate	5	7	8	4	4	19
Severe	1	2	3	2	1	7
Extremely severe	1	-	1	1	3	-
Anxiety						
Normal	7	10	5	10	6	26
Mild	3	1	4	5	4	9
Moderate	1	3	6	3	1	12
Severe	7	4	1	2	5	10
Extremely severe	3	4	4	4	5	10

Factors of Mental Health Problem

The factors contributing to mental health problems demonstrated clear gender differences (Table 3). For male students, family issues (76.2%) and financial issues (76.2%) were the most significant stressors. Self-related issues (66.7%) and past experiences (61.9%) were also important contributing factors.

In contrast, female students most frequently cited self-related issues (73.1%) and study-related stress (73.1%) as major factors. A high proportion of females also identified friends, colleagues, and peers as major contributors (71.6%). Furthermore, a greater proportion of female students (50.7%) identified physical appearance as an influencing factor compared to males.

Table 3: The factors of mental health problems among Medical Imaging students based on gender.

Factors	**Male, *n***	**Female, *n***
Family	16	42
Financial	16	36
Friends, Colleagues, Peers	11	48
Life events	12	41
Myself	14	49
Past experience	13	37
Physical appearance	9	34
Relationship	10	25
Self-expectation	12	44
Social media	4	16
Study	7	49
Trauma incidents	3	12
Others	1	3

Life Satisfaction and Self-Coping Approach

Life Satisfaction

Overall, the "Satisfaction of Life Test" showed a general sense of contentment, with a sizable proportion of respondents agreeing or partially agreeing with the positive statements (Table 4). However, the statement, "If I could live my life over, I would change almost nothing," elicited a polarised response, with 14 students strongly agreeing but 8 strongly disagreeing, suggesting that a significant portion of the cohort is critical of their life choices or circumstances.

Table 4: Satisfaction with Life Scale (SWLS) among Medical Imaging students.

Life satisfaction	**Number, *n***						
	Strongly agree	**Agree**	**Slightly agree**	**Neutral**	**Slightly disagree**	**Disagree**	**Strongly disagree**
In most ways my life is close to my ideal	5	22	25	25	7	3	1
The conditions of my life are excellent	5	27	18	28	8	1	1
I am satisfied with my life	10	25	20	20	8	5	0
So far, I have gotten the important things I want in my life	11	18	25	20	8	5	1
If I could live my life over, I would change almost nothing	14	12	15	19	11	9	8

Gender-based analysis of life satisfaction (Table 5) revealed that males consistently reported higher average satisfaction (e.g., mean score of 3.62 for "In most ways my life is close to my ideal") across all five statements compared to females (e.g., mean score of 3.10 for the same statement).

Table 5: The life satisfaction test among Medical Imaging students based on gender.

Life Satisfaction	Gender	Mean
In most ways my life is close to my ideal	Male	3.62
	Female	3.10
The conditions of my life are excellent	Male	3.24
	Female	3.13
I am satisfied with my life	Male	3.24
	Female	3.01
So far, I have gotten the important things I want in my life	Male	3.33
	Female	3.12
If I could live my life over, I would change almost nothing	Male	4.10
	Female	3.55

Self-Coping Approach

Self-coping strategies (Table 6) were primarily dominated by passive activities such as sleep and listen to music, and spiritual method for both genders. The most common coping mechanisms for both male and female students were sleeping (Male: 80.9%; Female: 80.6%), listening to music (Male: 76.2%; Female: 76.1%), and worshipping and praying to Allah SWT (Male: 71.4%; Female: 76.1%). Self-isolation was also a prevalent strategy among both genders, though slightly more common in females (62.7%) than males (57.1%). Females, relative to males, were more likely to engage in creative activities such as drawing, painting and reading. Conversely, males reported a higher percentage for physical activities such as exercising and sport.

Table 6: The self-coping approach to mental health problem Medical Imaging students based on gender.

Self-Coping Method	Gender			
	Male (N = 21)		Female (N = 67)	
	n	%	*n*	%
Browse social media	11	52.4%	36	53.7%
Cook/bake	3	14.3%	12	17.9%
Draw/painting	2	9.5%	19	28.4%
Eat	14	66.7%	47	70.1%
Exercise/sport	12	57.1%	14	20.9%
Explore new things	4	19.0%	12	17.9%
Go to entertainment centre	8	38.1%	16	23.9%
Hiking/outdoor activities	6	28.6%	10	14.9%
Listen to music	16	76.2%	51	76.1%
Photography	2	9.5%	4	6.0%
Play musical instrument	1	4.8%	5	7.5%
Read	0	0%	19	28.4%
Seek advice	7	33.3%	14	20.9%
Self-isolate	12	57.1%	42	62.7%
Share with trusted person	9	42.9%	31	46.3%
Shopping	5	23.8%	17	25.4%
Sleep	17	81.0%	54	80.6%
Spend time with family/friends	10	47.6%	22	32.8%
Travel	7	33.3%	9	13.4%
Watch drama/movie	12	57.1%	39	58.2%
Worship & pray to Allah SWT	15	71.4%	51	76.1%
Write poetry/fiction	1	4.8%	6	9.0%
Others	0	0%	0	0%

Overall, both male and female students use sleeping and listening to music as primary coping strategies. Religious practices like worship and prayer are equally important for both genders. Females, on the other hand, tend to engage in more social interactions and creative activities, whereas males prefer physical activities and solitary coping methods such as self-isolation. These findings suggest that mental health support programmes should take gender preferences into account in order to effectively address Medical Imaging students' coping needs. For example, providing opportunities for creative expression and social interaction may benefit female students, whereas encourage physical activity and provide spaces for quiet reflection may benefit male students.

DISCUSSION

Mental Health Status and Academic Stress

The study's findings regarding the Depression, Anxiety, and Stress Scale (DASS-21) indicate that a significant proportion of Medical Imaging students, particularly those in Year 1 to Year 3, experienced mild to extremely severe levels of mental distress. In contrast, Year 4 students reported the most stable mental health status because they are more academic experience and adaptation to collegiate demands and workload in learning. Anxiety was identified as the most prevalent issue across the entire cohort, with a concerning number of students exhibiting severe and extremely severe levels. This high prevalence

of anxiety warrants targeted intervention, particularly in the early academic years where students are adjusting to a new environment and curriculum.

Gender-Specific Stressors and Influencing Factors

The findings highlighted distinct patterns in the factors contributing to mental health challenges based on gender. Male students were primarily stressed by external factors, with family and financial issues being the most significant stressors. For female students, the major contributors were internal and social pressures, with academic performance (study-related stress), self-related issues, and social relationships with peers being the most highly cited factors. This gender-based variance emphasizes that mental health support programs must be tailored to address the unique stressors faced by each group, such as providing financial counselling for male students and establishing enhanced peer support networks for female students.

Life Satisfaction and Self-Coping Mechanisms

An analysis of life satisfaction indicated that male students reported a higher mean level of satisfaction across all related statements compared to female students. Regarding self-coping strategies, both genders heavily relied on passive and spiritual methods, with sleeping, listening to music, and engaging in religious practices being the most common. Gender differences in coping preferences were notable: males were more likely to utilize physical activities such as exercise and sports, along with self-isolation. Female students showed a greater inclination towards social strategies, such as discussing problems with trusted people, and creative outlets like reading and drawing. These insights are valuable for developing gender-sensitive intervention programs, such as encouraging physical activities for male students and facilitating creative expression and social support groups for female students.

CONCLUSION

The study concludes that a significant proportion of Medical Imaging students at IIUM face considerable mental health challenges, with anxiety emerging as the most prevalent concern. Crucially, the research identified distinct gender differences in both the primary factors influencing mental health and the preferred self-coping mechanisms. These findings underscore the necessity for university support systems to develop and implement gender-specific and targeted mental health interventions to effectively support the unique psychological needs of both male and female student populations. For instance, universities could introduce female-focused emotional expression workshops—such as guided art therapy or peer-support circles—since female students in this study showed higher engagement in activities like drawing, reading, and sharing with trusted persons. Meanwhile, male-focused stress-management programs could incorporate physical activity-based coping, such as structured sports sessions or outdoor challenges, aligning with the higher tendency of male students to cope through exercise, hiking, and social activities.

The study acknowledges limitations, including a small, non-random sample size, which restricts the generalizability of the findings. Data collection during an active semester may have inflated typical stress levels, and the reliance on self-reported data is susceptible to response biases. Future research should employ a larger, more diverse, and randomly selected sample, alongside longitudinal studies to monitor changes in mental health over time. The integration of qualitative research methods, such as interviews or focus groups, is also recommended to provide richer and more contextual insights into students' personal experiences, particularly given that the use of a single institution and self-report measures may introduce potential sources of bias in the current findings.

ACKNOWLEDGEMENT

This research was not funded by any grant.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author(s) used Gemini by Google to rephrase some sentences in the texts and formatting the references. The author(s) subsequently reviewed, revised, and approved all content, and accept full responsibility for the final manuscript.

REFERENCES

Hamza, C. A., Ewing, L., Heath, N. L., & Goldstein, A. (2021). When social isolation is nothing new: A longitudinal study on psychological distress during COVID-19 among university students with and without pre-existing mental health concerns. *Canadian Psychology/Psychologies Canadienne, 62(1), 20-30.* http://dx.doi.org/10.1037/cap0000255

Institute for Youth Research Malaysia, & United Nations Children's Fund. (2024). Malaysian Youth Mental Health Index 2023 (MyMHI'23). *Institute for Youth*

Research Malaysia & UNICEF Malaysia.

Jacka, F. N., Kremer, P. J., Berk, M., de Silva-Sanigorski, A. M., Moodie, M., Leslie, E. R., Pasco, J. A., & Swinburn, B. A. (2011). A prospective study of diet quality and mental health in adolescents. *PloS One, 6(9), e24805*. https://doi.org/10.1371/journal.pone.0024805

Jurewicz I. (2015). Mental health in young adults and adolescents - supporting general physicians to provide holistic care. *Clinical Medicine (London, England), 15(2), 151–154*. https://doi.org/10.7861/clinmedicine.15-2-151

Kadhum, M., Ayinde, O. O., Wilkes, C., Chumakov, E., Dahanayake, D., Ashrafi, A., Kafle, B., Lili, R., Farrell, S., Bhugra, D., & Molodysnki, A. (2022). Wellbeing, burnout and substance use amongst medical students: A summary of results from nine countries. *The International Journal of Social Psychiatry, 68(6), 1218–1222.* https://doi.org/10.1177/00207640221094955

Krejcie, R.V., & Morgan, D.W., (1970). Determining Sample Size for Research Activities. Educational and Psychological Measurement. Small-Sample Techniques (1960). *The NEA Research Bulletin, Vol. 38.*

Laranjeira, C., Querido, A., Sousa, P., & Dixe, M. A. (2023). Assessment and Psychometric Properties of the 21-Item Depression Anxiety Stress Scale (DASS-21) among Portuguese Higher Education Students during the COVID-19 Pandemic. *European Journal of Investigation in Health, Psychology and Education, 13(11), 2546–2560*. https://doi.org/10.3390/ejihpe13110177

Liu, Z., Xie, Y., Sun, Z., Liu, D., Yin, H., & Shi, L. (2023). Factors associated with academic burnout and its prevalence among university students: a cross-sectional study. *BMC Medical Education, 23(1), 317*. https://doi.org/10.1186/s12909-023-04316-y

Milojevich, H. M., & Lukowski, A. F. (2016). Sleep and Mental Health in Undergraduate Students with Generally Healthy Sleep Habits. *PloS One, 11(6), e0156372.* https://doi.org/10.1371/journal.pone.0156372

Pascoe, M. C., Hetrick, S. E., & Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth, 25(1), 104-112*. https://doi.org/10.1080/02673843.2019.1596823

Sandhu, S. S., Ismail, N. H., & Rampal, K. G. (2015). The Malay Version of the Perceived Stress Scale (PSS)-10 is a Reliable and Valid Measure for Stress among Nurses in Malaysia. *The Malaysian Journal of Medical Sciences : MJMS, 22(6), 26–31*.

Slavin, S. J., Schindler, D. L., & Chibnall, J. T. (2014). Medical student mental health 3.0: improving student wellness through curricular changes. *Academic medicine : Journal of the Association of American Medical Colleges, 89(4), 573–577*. https://doi.org/10.1097/ACM.0000000000000166

Sulaiman, N. A. (2022, October 10). Masalah kesihatan mental membimbangkan – Khairy. *Berita Harian Online.* https://www.bharian.com.my/berita/nasional/2022/10/1010651/masalah-kesihatan-mental-membimbangkan-khairy (Retrieved: 12 December 2025).

The Star. (2024, April 22). Malaysian youths' mental health status at moderate level, study finds. *The Star*. https://www.thestar.com.my/news/nation/2024/04/22/malaysian-youths039-mental-health-status-at-moderate-level-study-finds (Retrieved December 12, 2025)

Townshend, K. (2025). Satisfaction with Life Scale (SWLS). In: Medvedev, O.N., Krägeloh, C.U., Siegert, R.J., Singh, N.N. (eds) Handbook of Assessment in Mindfulness Research. *Springer, Cham*. https://doi.org/10.1007/978-3-031-47219-0_83

Yaghmour, K. A., Alhmyri, S. M., Alhmyri, B. M., Sharaf, R., Alasmari, M. A., & Almilabi, M. M. (2023). The Effect of Sleep Quality and Mental Health on Academic Performance Among the Medical Students of King Abdulaziz University. *Cureus, 15(9), e44951*. https://doi.org/10.7759/cureus.44951

Yang, C., Chen, A., & Chen, Y. (2021). College students' stress and health in the COVID-19 pandemic: The role of academic workload, separation from school, and fears of contagion. *PloS One, 16(2), e0246676*. https://doi.org/10.1371/journal.pone.0246676

ORIGINAL ARTICLE

Utilising MRI Data and VOI-Based Analysis for Enhanced Epilepsy Prediction: A Translational Approach from Bench to Bedside

Nur Arif Che Mat[1], Nur Nadiah Mohd Rais[2], Mohd Zulfaezal Che Azemin[3], Intan Bazilah Abu Bakar[4], Iqbal Jamaludin[2,5*]

[1]Radiology Department, Hospital Tengku Ampuan Afzan, 25100 Kuantan, Pahang, Malaysia
[2]Department of Diagnostic Imaging and Radiotherapy, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia
[3]Department of Optometry and Visual Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia
[4]Department of Radiology, Kulliyyah of Medicine, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia
[5]Integrated Omics Research Group (IORG), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
MRI; VOI-Based Analysis; Epilepsy Prediction; Brain

ABSTRACT

Background: Epilepsy is one of the most prevalent neurological disorders globally, profoundly impacting patient's quality of life and stretching healthcare resources. Despite technological advances in neuroimaging, early and accurate detection of epileptic foci remains elusive, especially when standard MRI scans appear structurally normal. For clinicians, radiologists and neurologists, the limitations of subjective interpretation underscore an urgent need for objective diagnostic methods. This study addresses that gap by introducing Volume of Interest (VOI)-based analysis as an innovative tool to detect microstructural brain abnormalities associated with epilepsy. **Methods:** This study involves a retrospective analysis of 30 brain MRI images (T1-Weighted MPRAGE), comprising of 12 epilepsy patients and 18 matched controls (11 males, 19 females; mean age 42.7 ± 17.2 years). The images were converted from the DICOM format into the Brain Imaging Data Structures (BIDS) standard using the fMRIPrep platform, which were then normalised following the standard Montreal Neurological Institute (MNI) template. The images were segmented into Grey Matter, White Matter and Cerebrospinal Fluid. The segmented Grey Matter regions have then been analysed using VOI-based analysis from the Harvard-Oxford Cortical Structural Atlas. The VOI-based analysis results were then statistically tested using ANCOVA with False Detection Rate (FDR) correction. **Results:** Three regions show statistically significant ANCOVA results, which were Superior Frontal Gyrus (p=0.041), Superior Parietal Lobule (p=0.026) and Lingual Gyrus (p=0.036). However, all three regions fail the FDR correction (q=0.36). **Conclusion:** This work shows that combining MRI data with VOI-based analysis can reveal subtle structural patterns in the superior frontal gyrus, superior parietal lobule and lingual gyrus that may contribute to the understanding of epilepsy risk. Although these patterns were not sustained after FDR correction, they offer a promising direction for more objective predictive tools. Continued research with larger cohorts is essential to confirm these early signals and strengthen their relevance for epilepsy prediction.

INTRODUCTION

Epilepsy is a long-standing neurological condition marked by recurrent, unprovoked seizures that can affect many aspects of a person's life. With an estimated 50 million people affected globally, the disorder continues to place a significant burden on individuals, families and healthcare systems (World Health Organisation, 2023; Beghi, 2020). In Malaysia, the prevalence has been estimated at 7.8 cases per 1,000 people, reflecting a substantial public health concern (Fong et al., 2021). Certain groups such as younger adults, males and individuals with focal epilepsy or poorly controlled seizures remain at higher risk of complications, including Sudden Unexpected Death in Epilepsy (SUDEP) (Khor et al, 2021; 2022). These factors highlight the continuing need to improve the tools used for early detection and accurate diagnosis.

Magnetic Resonance Imaging (MRI) plays a central role in evaluating epilepsy, providing essential structural information that informs diagnosis and treatment planning. While advanced MRI technologies hold great potential, clinical interpretation still relies heavily on qualitative assessment, and subtle abnormalities can easily go unnoticed (Bernasconi et al., 2019).

Recent studies have shown that combining structural MRI with computational analysis can markedly improve diagnostic accuracy; for example, machine learning approaches such as Support Vector Machines (SVM) have achieved up to 94% accuracy in distinguishing epileptic from non-epileptic scans (Jamaludin al., 2022). These

* *Corresponding author.*
E-mail address: iqbaljamaludin@iium.edu.my

findings suggest that data-driven methods may help capture features that traditional visual diagnosis may miss, particularly in focal epilepsy where small structural variations may influence treatment decisions (Caldairou et al., 2022; Bank et al., 2023). Fractal analysis has gained attention in neuroimaging for its ability to quantify structural complexity within the brain. By estimating Fractal Dimension (FD), researchers can detect subtle changes in grey matter structures that may not be apparent on the routine scans (Gleichgerrcht et al., 2021; Ziukelis et al., 2022). This type of quantitative assessment offers an objective way to characterise brain morphology and has shown promise in various neurological conditions.

More recently, local or regional assessment such as Volume of Interest (VOI)-based analysis, has been used to identify fine-detailed structural differences across specific brain regions. Studies have demonstrated the effectiveness of this approach in detecting intricate neuroanatomical variations linked to cognitive and neurological traits (Jamaludin et al., 2022). Applying VOI-based analysis to epilepsy may therefore offer a powerful extension to conventional MRI interpretation by capturing regional patterns associated with the condition.

Although quantitative imaging methods are gaining traction, the potential of VOI-based analysis for predicting epilepsy remains largely unexplored. Current diagnostic practices still depend heavily on subjective MRI interpretation, which may overlook the subtle microstructural changes that characterise epileptogenic brain tissues. Bridging this gap is important for improving diagnostic reliability and for advancing objective, image-based markers of epilepsy. In response, this study aimed to determine whether an automated, atlas-based VOI volumetric pipeline applied to high-resolution T1-weighted MRI of brain can discriminate patients with epilepsy from healthy controls. Through these objectives, the study seeks to contribute to a more rigorous and objective approach to epilepsy prediction using structural MRI.

MATERIALS AND METHODS

Study Design and Data Overview

This study was conducted using a retrospective design involving structural brain MRI scans obtained from the Department of Radiology, Sultan Ahmad Shah Medical Centre (SASMEC @IIUM). Three ethical clearances were obtained; 1) Kulliyyah Postgraduate and Research Committee (KPGRC) (ID Number: KAHS 40/25), 2) IIUM Research Ethical Committee (IREC) (IREC Number: IREC 2025-086) and 3) SASMEC @IIUM Research Committee (SASRC) (Reference Number: IIUM/423/013/14/1/SRG25-072-0072).

MRI Acquisition

The MRI system used was 3 Tesla Siemens Magnetom Spectra Scanner (Siemens Medical Solutions, Erlangen, Germany). Its gradient strength is 33 mT/m with slew rate of 125 T/m/s. The standard T1-Weighted 3D-MPRAGE (3D-Magnetisation Prepared Rapid Gradient Echo) protocol was used, with TR value of 1880 ms, TE of 3 ms and TA of 4.23 minute. The voxel size of 1.0 mm x 1.0 mm x 1.0 mm was used, with the FoV in the read direction of 250 mm. The slice thickness during acquisition was set at 45.5% where the phase encoding was set from right part of the brain to the left direction. The number of slices per image slab was 176. The raw imaging data were originally stored in native Digital Imaging and Communication in Medicine (DICOM) format within clinical archives.

A total of thirty T1-weighted MPRAGE brain MRI images from the scans conducted between the year 2020 until 2024 were retrieved, representing twelve patients clinically diagnosed with epilepsy (confirmed by EEG test results) and eighteen matched healthy controls (11 males, 19 females; mean age 42.7 ± 17.2 years). The inclusion criteria are epilepsy patients diagnosed based on clinical history, EEG, and MRI results for epilepsy group, while healthy controls were confirmed with no history of neurological disorders by radiologist and neurologist. Age for both groups were 18 to 60 years old, for both male and female. Patients with comorbid neurological conditions (e.g., stroke, dementia). MRI contraindications (e.g., metal implants) and active pregnancy were excluded from this study. Both groups were matched for age and gender. The sample size was determined using the Power and Sample Size Program, guided by parameters published by Chang et al. (2023), and all eligible cases in the archive were used for this retrospective study.

Preprocessing Pipeline

Image conversion and preparation

The raw imaging data were originally stored in native Digital Imaging and Communication in Medicine (DICOM) format within clinical archives. To ensure a uniform data structure suitable for automated neuroimaging pipelines, the scans were converted into the Brain Imaging Data Structure (BIDS) standard. This conversion was facilitated by a custom workflow built around the script *dcm_to_mni.py*, which systematically detected DICOM series for each subject and executed conversion using *heudiconv*. The script first attempted the standard *reproin* heuristic to identify the T1-weighted anatomical series. In

instances where this approach did not create BIDS-compliant directory structure, a fallback custom heuristic routed the MPRAGE sequence into the appropriate anatomical folder within the BIDS hierarchy. The resulting dataset conformed to the BIDS specification, with each participant assigned a unique subject identifier located within a cleanly organised directory structure.

Spatial Normalisation

All BIDS datasets were subsequently processed using the fMRIPrep platform, with additional preprocessing steps orchestrated via *dcm_to_mni.py*. Each subject's T1-weighted image was skull-stripped, bias-field corrected, and spatially normalised to the MNI152NLin2009cAsym template at 1-mm isotropic resolution. fMRIPrep also produced tissue probability maps for grey matter, white matter and cerebrospinal fluid. These outputs formed the foundation for the study's volumetric analysis, which focused on the extraction of both global tissue estimates and region-specific grey-matter volumes.

Tissue segmentation

Tissue segmentation was performed automatically through fMRIPrep's anatomical workflow. Grey matter, white matter and cerebrospinal fluid volumes were estimated by summing voxel-wise tissue probabilities, allowing the computation of global tissue measures expressed in millilitres. Total intracranial volume was estimated as the sum of the grey matter, white matter and cerebrospinal fluid; and was later included as a covariate in the group-comparison analysis.

VOI-Based Regional Analysis

Regional morphometric analysis was conducted using a VOI-based framework implemented in *analyse_VOI_HO.py*. This process involved integrating the Harvard-Oxford Cortical Structural Atlas with the grey-matter probability maps produced by fMRIPrep. For each subject, the atlas was resampled to match the individual's grey-matter anatomical space, ensuring spatial correspondence between atlas labels and tissue probabilities. Regional volumes were calculated by applying atlas-defined masks to the grey-matter probability maps and converting the resulting weighted sums into millilitres measurements. The script generated a consolidated dataset containing demographic information, global tissue measurements and regional VOI-based grey-matter volumes for all subjects.

This automated framework supports a repeatable, objective, and clinically transferable neuroimaging pipeline, integrating standardised conversion, robust preprocessing, and rigorous analysis to advance MRI-based biomarkers for more accurate epilepsy diagnosis.

Statistical Analysis

To assess neuroanatomical differences between epilepsy and control groups, an Analysis of Covariance (ANCOVA) was performed for each global and regional volume, using Statistical Package for the Social Sciences (SPSS) version 29. The models incorporated age and intracranial volume as covariates to account for inherent biological variability. Group differences were evaluated using ANOVA-derived F statistics extracted from each model. Given the large number of atlas-derived regions analysed, the Benjamini-Hochberg False Discovery Rate (FDR) correction was applied to adjust for multiple comparisons. Regions meeting an adjusted significance threshold of $q<0.05$ were considered to show statistically robust volumetric differences between groups.

This study presents a fully automated structural MRI pipeline designed to characterise epilepsy-related neuroanatomical differences from T1-weighted images (Figure 1). The workflow standardises preprocessing and segmentation to generate MNI-normalised images and tissue probability maps, enabling consistent extraction of atlas-based regional grey-matter volumes. These volumetric biomarkers are then compared between groups using ANCOVA with FDR correction. Overall, the pipeline supports transparent, reproducible analysis with potential for translation into routine neuroimaging workflows.

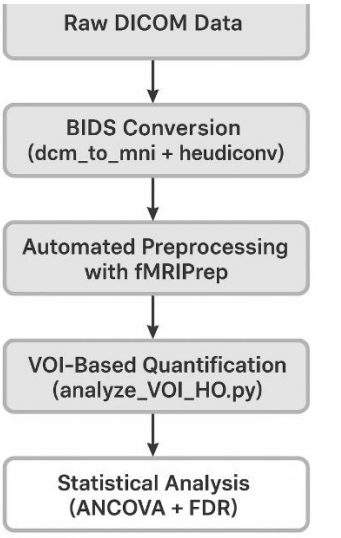


Figure 1: Automated T1-weighted MRI workflow: DICOM→BIDS (*heudiconv*), preprocessing with *fMRIPrep* (MNI-space outputs, tissue maps), Harvard–Oxford atlas grey-matter volumetry, and group comparison using ANCOVA with FDR correction.

RESULTS

Demographic Characteristics and Data Pre-Processing

This study used retrospective MRI brain images data from 30 participants, comprising of 12 individuals diagnosed with epilepsy and 18 age- and gender-matched healthy control subjects. The overall demographic profile indicated a mean age of 42.7 ± 17.2 years and a gender distribution of 11 males and 19 females.

High-resolution T1-weighted MPRAGE structural MRI sequences were acquired and utilised for all subsequent morphometric analysis. Data preparation involved a rigorous and reproducible pipeline: images were first converted from the DICOM format to the BIDS standard using the fMRIPrep platform. These data were subsequently normalised to the standard MNI template space. Following normalisation, images were meticulously segmented into grey matter, white matter and cerebrospinal fluid. Grey matter volume (GMV) were then derived using a VOI-based analysis across all 48 predefined cortical regions outlined by the Harvard-Oxford Cortical Structural Atlas.

Analysis of Covariance (ANCOVA) Findings (Uncorrected)

ANCOVA was used to statistically evaluate the inter-group differences in regional grey matter volume, adjusting for age and total intracranial volume (TIV) as covariates to mitigate potential confounding effects. Gender was excluded from the covariates as the images were retrieved retrospectively, and the limitation in getting the large number of images during the data collection process.

The initial inferential analysis performed using an uncorrected significance threshold of $p<0.05$, identified three specific cortical regions that exhibited a statistically significant volume alteration in the epilepsy group compared to controls. In all three regions, the epilepsy group demonstrated a reduced mean grey matter volume. The statistical details of these uncorrected findings are summarised in Table 1.

Table 1: Cortical VOIs showing statistically significant differences in grey matter volume between the epilepsy and control group using an uncorrected threshold ($p<0.05$).

Region	Mean (Epilepsy)	Mean (Control)	Difference	p-value
Superior Parietal Lobule	0.6791	0.6791	0.621	0.026
Lingual Gyrus	0.3798	0.3798	0.267	0.036
Superior Frontal Gyrus	0.2326	0.2326	0.093	0.041

Correction for Multiple Comparisons

Given the high number of concurrent statistical tests (n=48) performed across the anatomical atlas, the False Discovery Rate (FDR) method was applied to the entire p-value distribution to control the expected proportion of false positive findings (Type 1 errors).

Critically, following FDR control for multiple comparisons, no cortical VOI maintained a statistically significant difference at the controlled threshold of $q < 0.05$. Table 2 presents the comparison between the uncorrected p-values and the resulting FDR-corrected 1-values for the three regions identified in the initial analysis.

Table 2: q=FDR-corrected p-value for regions identified in the uncorrected ANCOVA analysis. No region met the significance threshold after controlling the False Discovery Rate across all 48 VOI tests.

Region	Corrected q-value (*FDR)	Partial η^2	**95% CI for Partial η^2
Superior Parietal Lobule	0.36	0.089	0.000 to > 1.0
Lingual Gyrus	0.48	0.094	0.000 to > 1.0
Superior Frontal Gyrus	0.51	0.106	0.000 to > 1.0

*FDR: False Discovery Rate

** Note: Upper confidence limits for Partial η^2 are unbounded due to the properties of the noncentral F distribution at moderate F values and sample size.

The comparison of the VOI-based values between the epilepsy and control groups showed small but detectable effects across the three regions of interest (ROIs). For the Superior Parietal Lobule, the group difference yielded $F(1,48) = 4.6938$ with a corresponding partial η^2 of 0.089. For the Lingual Gyrus, the analysis produced $F(1,48) = 4.9816$ and a partial η^2 of 0.094. For the Superior Frontal Gyrus, the observed effect was $F(1,48) = 5.7179$, reflecting a partial η^2 of 0.106.

The 95% confidence intervals (CIs) for partial η^2 showed lower bounds approximating 0.000 for all three ROIs, while the upper bounds were unbounded. This is a known consequence of inverting the noncentral F distribution in situations with modest F values and moderate sample sizes. The unbounded upper interval does not represent a computational error but reflects the limited precision with which the true effect size can be estimated. These findings collectively indicate the presence of small but meaningful regional differences despite the modest effect magnitudes.

This result confirms that while the three identified regions demonstrated a suggestive but not statistically robust after correction, the observed effect sizes were insufficient to pass the critical threshold necessary to declare a robust,

statistically independent difference when controlling for the large number of simultaneous comparisons.

DISCUSSION

This study set out to explore whether a fully automated, VOI-based morphometric analysis of high-resolution T1-weighted (MPRAGE) MRI could reliably detect subtle regional grey matter volume (GMV) differences between adults with epilepsy and matched healthy controls. While initial uncorrected analysis suggested reduced GMV in three cortical regions among epilepsy patients, these findings did not survive correction for multiple comparisons, highlighting both the promise and the current limitations of this approach. Thus, we did not identify robust volumetric biomarkers of epilepsy using this VOI-based approach in this modest sample.

The challenge of identifying robust, statistically significant regional brain abnormalities in epilepsy, especially in MRI-negative or nonlesional cases, is well documented in the literature. Recent studies have shown that advanced morphometric techniques, such as voxel-based and surface-based morphometry, can reveal subtle structural changes in both focal and generalised epilepsy, but their sensitivity and specificity are highly variable and often depend on sample size, imaging quality and analytic rigour (Bunyamin et al., 2025; Feng et al., 2025; Di Micco et al., 2025; Yang et al., 2024). For example, a large-scale study using the morphometric analysis program (MAP) in truly MRI-negative focal epilepsy found that, while MAP could detect previously overlooked lesions in a minority of cases, the overall diagnostic yield was limited, with a high rate of false positives and modest balanced accuracy (51-60%)(Feng et al., 2025). Similarly, a recent review emphasised that, although post-processing methods can convert some MRI-negative cases to MRI-positive, the concordance with true epileptogenic zones remains inconsistent, and integration with electrophysiological data is often necessary for clinical decision-making (Bunyamin et al., 2025).

The present findings are consistent with these reports, underscoring the complexity of epilepsy as a network disorder. Structural changes may be subtle, distributed, or below the threshold of detection for current morphometric pipelines, especially in small or heterogeneous groups. Notably, studies in temporal lobe epilepsy and idiopathic generalised epilepsy have demonstrated that GMV alterations often extend beyond the primary epileptogenic focus, involving widespread cortical and subcortical networks (Huang et al., 2025; Jber et al., 2021). However, the effect sizes are frequently modest, and robust group differences may only emerge in larger, multi-centre studies or when combined with other imaging modalities such as PET or advanced radiomics (Zhang et al., 2025; Tsai et al., 2025).

Recent advances in machine learning and radiomics have shown promise in improving the detection and prediction of epilepsy-related brain changes, particularly when integrating multiple imaging features and clinical data (Tsai et al., 2025; Yang et al., 2024). For instance, automated ROI-based morphometric features combined with machine learning have achieved high accuracy in distinguishing temporal lobe epilepsy from controls, even in MRI-negative cases (Yang et al., 2024). However, these approaches require large, well-annotated datasets and careful validation to avoid overfitting and ensure generalisability.

CONCLUSION

In summary, this study demonstrates the feasibility and reproducibility of a fully automated, VOI-based morphometric MRI analysis pipeline for investigating regional brain morphology in epilepsy. While initial findings suggested trends toward reduced grey matter volume in specific cortical regions among epilepsy patients, these differences did not reach statistical significance after correction for multiple comparisons. Consistent with current literatures, these findings are best interpreted as a methodological contribution, demonstrating the practical implementation and application of a fully automated, reproducible volumetric pipeline within a retrospective epilepsy cohort. (Bunyamin et al., 2025; Fearns et al., 2023; Di Micco et al., 2025; Yang et al., 2024).

The results underscore the need for larger, multi-centre studies, integration of multi-modal imaging (e.g., PET/MRI, radiomics), and the development of advanced analytic techniques to enhance sensitivity and specificity. Future research should also consider combining structural imaging with electrophysiological and clinical data to improve the localisation and characterisation of epileptogenic zones (Bunyamin et al., 2025; Feng et al., 2025; Yang et al., 2024). Ultimately, while VOI-based morphometric analysis holds promise as an objective, reproducible tool for epilepsy research and potentially for clinical practice, its current diagnostic utility remains limited in small, heterogeneous groups. Ongoing methodological refinement and collaborative research will be essential to realise its full potential in the early and accurate detection of epilepsy-related brain abnormalities.

ACKNOWLEDGEMENT

This research was funded by the SASMEC Research Grant (Project ID: SRG25-072-0072).

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

Artificial intelligence was used to improve content development and editing in the preparation of this manuscript. The authors subsequently reviewed, revised and approved all content and accept all responsibility for the final manuscript, and the authors guarantee its integrity and accuracy.

REFERENCES

Bank, A. M., Kuzniecky, R., Knowlton, R. C., Cascino, G. D., Jackson, G., Pardoe, H. R., & Human Epilepsy Project Investigators (2022). Structural Neuroimaging in Adults and Adolescents With Newly Diagnosed Focal Epilepsy: The Human Epilepsy Project. *Neurology*, *99*(19), e2181–e2187. https://doi.org/10.1212/WNL.0000000000201125

Beghi E. (2020). The Epidemiology of Epilepsy. *Neuroepidemiology*, *54*(2),185–191. https://doi.org/10.1159/000503831

Bernasconi, A., Cendes, F., Theodore, W. H., Gill, R. S., Koepp, M. J., Hogan, R. E., Jackson, G. D., Federico, P., Labate, A., Vaudano, A. E., Blümcke, I., Ryvlin, P., & Bernasconi, N. (2019). Recommendations for the use of structural magnetic resonance imaging in the care of patients with epilepsy: A consensus report from the International League Against Epilepsy Neuroimaging Task Force. *Epilepsia*, *60*(6),1054–1068. https://doi.org/10.1111/epi.15612

Bunyamin, J., Sinclair, B., Law, M., Kwan, P., O'Brien, T. J., & Neal, A. (2025). Voxel-based and surface-based cortical morphometric MRI applications for identifying the epileptogenic zone: A narrative review. *Epilepsia open*, *10*(2),380–397. https://doi.org/10.1002/epi4.70012

Caldairou, B., Foit, N. A., Mutti, C., Fadaie, F., Gill, R., Lee, H. M., Demerath, T., Urbach, H., Schulze-Bonhage, A., Bernasconi, A., & Bernasconi, N. (2021). MRI-Based Machine Learning Prediction Framework to Lateralize Hippocampal Sclerosis in Patients With Temporal Lobe Epilepsy. *Neurology*, *97*(16), e1583–e1593. https://doi.org/10.1212/WNL.0000000000012699

Chang, A. J., Roth, R., Bougioukli, E., Ruber, T., Keller, S. S., Drane, D. L., Gross, R. E., Welsh, J., Abrol, A., Calhoun, V., Karakis, I., Kaestner, E., Weber, B., McDonald, C., Gleichgerrcht, E., Bonilha, L., & Alzheimer's Disease Neuroimaging Initiative (2023). MRI-based deep learning can discriminate between temporal lobe epilepsy, Alzheimer's disease, and healthy controls. *Communications medicine*, *3*(1), 33. https://doi.org/10.1038/s43856-023-00262-4

Di Micco, V., Mercier, M., Espagnet, C., Napolitano, A., Benedictis, A., Carfì-Pavia, G., Marras, C., Vigevano, F., De Palma, L., & Specchio, N. (2025). Diagnostic performance of morphometric analysis program in pediatric epilepsy surgery. *Epilepsy & Behavior*, 171. https://doi.org/10.1016/j.yebeh.2025.110566

Fearns, N., Birk, D., Bartkiewicz, J., Rémi, J., Noachtar, S., & Vollmar, C. (2023). Quantitative analysis of the morphometric analysis program MAP in patients with truly MRI-negative focal epilepsy. *Epilepsy Research*, 192. https://doi.org/10.1016/j.eplepsyres.2023.107133

Feng, T., Yang, Y., Wang, Y., Fan, X., Wang, T., Chen, S., Wei, P., Shan, Y., & Zhao, G. (2025). Mapping grey matter and network abnormalities in seizure-onset patterns via stereotactic EEG. *Neurobiology of Disease*, 213. https://doi.org/10.1016/j.nbd.2025.107021

Fong, S. L., Lim, K. S., Tan, L., Zainuddin, N. H., Ho, J. H., Chia, Z. J., Choo, W. Y., Puvanarajah, S. D., Chinnasami, S., Tee, S. K., Raymond, A. A., Law, W. C., & Tan, C. T. (2021). Prevalence study of epilepsy in Malaysia. Epilepsy research, 170, 106551. https://doi.org/10.1016/j.eplepsyres.2021.106551

Gleichgerrcht, E., Munsell, B. C., Alhusaini, S., Alvim, M. K. M., Bargalló, N., Bender, B., Bernasconi, A., Bernasconi, N., Bernhardt, B., Blackmon, K., Caligiuri, M. E., Cendes, F., Concha, L., Desmond, P. M., Devinsky, O., Doherty, C. P., Domin, M., Duncan, J. S., Focke, N. K., Gambardella, A. ENIGMA-Epilepsy Working Group (2021). Artificial intelligence for classification of temporal lobe epilepsy with ROI-level MRI data: A worldwide ENIGMA-Epilepsy study. *NeuroImage. Clinical*, 31, 102765. https://doi.org/10.1016/j.nicl.2021.102765

Huang, X., Mao, L., Huang, K., Li, Y., Liu, P., Cao, D., Li, W., Li, W., Zhao, L., Zhang, Y., & An, D. (2025). Voxel-based meta-analysis of gray matter abnormalities in idiopathic generalized epilepsy. *Cerebral cortex*, 35 8. https://doi.org/10.1093/cercor/bhaf201

Jamaludin, I., Che Azemin, M. Z., Mohd Tamrin, M. I., &

Sapuan, A. H. (2022). Volume of Interest-Based Fractal Analysis of Huffaz's Brain. *Fractal Fract*, 6(7), 396. https://doi.org/10.3390/fractalfract6070396

Jber, M., Habibabadi, J., Sharifpour, R., Marzbani, H., Hassanpour, M., Seyfi, M., Mobarakeh, N., Keihani, A., Hashemi-Fesharaki, S., Ay, M., & Nazem-Zadeh, M. (2021). Temporal and extratemporal atrophic manifestation of temporal lobe epilepsy using voxel-based morphometry and corticometry: clinical application in lateralization of epileptogenic zone. *Neurological Sciences*, 42, 3305 - 3325. https://doi.org/10.1007/s10072-020-05003-2

Khor, S. B., Lim, K. S., Fong, S. L., Ho, J. H., Koh, M. Y., & Tan, C. T. (2021). Mortality in adult epilepsy patients in Malaysia: a hospital-based study. *Seizure,* 88, 56–59. https://doi.org/10.1016/j.eplepsyres.2022.106887

Khor, S. B., Lim, K. S., Fong, S. L., Ho, J. H., Koh, M. Y., & Tan, C. T. (2022). Cause of mortality among people with epilepsy in Malaysia: A hospital-based study. *Epilepsy research*, 181, 106887.

Tsai, M., Hsieh, K., Liu, Y., Yang, Y., Chang, H., Wong, T., & Peng, S. (2025). Morphometric and radiomics analysis toward the prediction of epilepsy associated with supratentorial low-grade glioma in children. *Cancer Imaging*, 25. https://doi.org/10.1186/s40644-025-00881-1

World Health Organisation (2023). Epilepsy. Retrieved November 8, 2023, from https://www.who.int/news-rro/fact-sheets/detail/epilepsy

Yang, L., Peng, B., Gao, W., A, R., Liu, Y., Liang, J., Zhu, M., Hu, H., Lu, Z., Pang, C., Dai, Y., & Sun, Y. (2024). Automated detection of MRI-negative temporal lobe epilepsy with ROI-based morphometric features and machine learning. *Frontiers in Neurology*, 15. https://doi.org/10.3389/fneur.2024.1323623

Zhang, C., Wang, Z., Wang, Y., Cao, H., Ren, L., Yu, T., Shan, Y., He, X., Duncan, J., Lu, J., Wei, P., & Zhao, G. (2025). Quantitative Cortex-Based Mapping With Hybrid 18F-FDG-PET/MR Images in MRI-Negative Epilepsy. *CNS Neuroscience & Therapeutics*, 31. https://doi.org/10.1111/cns.70336

Ziukelis, E. T., Mak, E., Dounavi, M. E., Su, L., & T O'Brien, J. (2022). Fractal dimension of the brain in neurodegenerative disease and dementia: A systematic review. *Ageing research reviews*, 79, 101651. https://doi.org/10.1016/j.arr.2022.101651

ORIGINAL ARTICLE

Investigating the Association Between Body Mass Index (BMI) and Body Image Perceptions Among High School Students

Nur Fadhlin Najwa Rosman[1], Roszanadia Rusali[1,2*] Aliza Haslinda Hamirudin[1,2], Ibrahim Abu Bakar[1,2]

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia
[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
body mass index (BMI); body image perception; adolescents; Malaysia

ABSTRACT

Background: Adolescence is a critical period for physical growth and psychological development, during which body mass index (BMI) and body image perceptions play an important role in health and well-being. Concerns about overweight, obesity, and body dissatisfaction are increasingly common among Malaysian adolescents. Therefore, this study aimed to determine the association between BMI status and body image perceptions among high school students in Kuantan, Pahang. **Methods:** A cross-sectional study was conducted among 140 students aged 13 to 15 years in Kuantan, Pahang. Participants were selected using convenience sampling, and only those with parental consent and valid responses were included. Anthropometric measurements were used to calculate BMI, while body image perceptions were assessed using a validated questionnaire, Body Self-Image Questionnaire Short-Form (BSIQ-SF), across four domains: Negative Affect, Attractiveness Evaluation, Physical Functionality Awareness, and Height Dissatisfaction. Statistical analyses included descriptive statistics, independent t-tests, and Pearson correlation. **Results:** The majority of students (75.0%) had a normal BMI, while 10.0% were overweight, 7.1% were obese, 4.3% were thin, and 3.6% were severely thin. Female students reported significantly greater dissatisfaction in the Height Dissatisfaction domain (p = 0.031). No significant gender differences were found in the Negative Affect, Attractiveness Evaluation, and Physical Functionality Awareness domains. BMI was significantly correlated with Negative Affect (r = 0.338, p < 0.001) and Physical Functionality Awareness (r = 0.218, p = 0.010), suggesting that students with higher BMI were more likely to experience body dissatisfaction and greater awareness of physical functionality. **Conclusion:** Our findings reveal that BMI and gender significantly influence body image perceptions among adolescents, reflecting both overweight/obesity and underweight concerns. Despite the growing importance of adolescent body image in public health, there remains a need for further research on sociocultural and environmental factors shaping these perceptions.

INTRODUCTION

Adolescence is a critical developmental stage characterized by rapid physical, psychological, and social changes that shape self-identity and self-perception (National Academies of Sciences, Engineering, and Medicine [NASEM], 2019) During this period, body changes associated with puberty often interact with sociocultural pressures to achieve idealised standards, contributing to heightened body image concerns (Khor et al., 2009; McCabe et al., 2003). In Malaysia, these concerns are widespread, with nearly 80% of adolescents reporting dissatisfaction with their body size or shape (Farah Wahida et al., 2011; Khor et al., 2009).

* *Corresponding author.*
E-mail address: roszanadia@iium.edu.my

Body weight status, usually measured with BMI, is closely connected to body image perceptions. Malaysia faces a double burden of malnutrition, with adolescents experiencing both undernutrition and rising rates of overweight and obesity. In Pahang, 31% of adolescents were classified as overweight or obese, while others fell into thinness categories (Jamani et al., 2020; NHMS, 2022). Misperceptions about body weight are common, especially among those aiming to be thinner, and are linked to unhealthy behaviors like dieting, binge eating, and physical inactivity (Pon et al., 2004; Toselli et al., 2023). Gender differences also influence body image, with females more dissatisfied with slimness and males aiming for muscularity (Avci & Keven Akliman, 2018), while social media and peer influence increase these pressures (Holland & Tiggemann, 2016).

Despite these concerns, limited recent evidence exists on the relationship between BMI and body image perceptions among Malaysian adolescents, particularly at

the high school level. Existing studies are either outdated or conducted among older populations, leaving a gap in understanding the current status of younger adolescents in Kuantan, Pahang. To address this gap, the present study examined the association between BMI status and body image perception among high school students in Kuantan, Pahang.

MATERIALS AND METHODS

Study Design and Setting

This cross-sectional study was conducted among high school students aged 13–16 years in Kuantan, Pahang. The school was purposively selected due to its large and diverse student population, considered representative of adolescents in the district.

Sample Size and Sampling

The sample size was calculated using a single proportion f formula, based on a previously reported prevalence of body dissatisfaction (Farah Wahida et al., 2011). The participants were recruited using convenience sampling. The inclusion criteria were: (i) male and female students aged 13–16 years, (ii) proficiency in Bahasa Melayu and English, and (iii) enrolment in the selected school. Students with chronic illnesses such as diabetes, heart disease, and cancer, or living in boarding facilities, were excluded.

Ethical Considerations

Ethical approval was obtained from the IIUM Research Ethics Committee (IREC No.: IREC 2025-KAHS/DNS16), the Ministry of Education, the Pahang State Education Department, and the school principal. Written informed consent was obtained from parents or legal guardians before participation.

Data Collection

Data collection was conducted during school hours using a structured questionnaire administered face-to-face. The instrument consisted of three parts: (i) sociodemographic information, (ii) anthropometric measurements, and (iii) body image perceptions. Height and weight were measured using standardised procedures with portable stadiometers and digital scales. BMI was calculated and classified according to the WHO's BMI-for-Age (z-score) classification for both genders (WHO, 2007). Body image perception was assessed using the validated Malay version of the Body Self-Image Questionnaire-Short Form (BSIQ-SF), comprising four domains: Negative Affect, Attractiveness Evaluation, Physical Functionality Awareness, and Height Dissatisfaction. Negative Affect captured negative emotions related to body image, while Attractiveness Evaluation reflected satisfaction with appearance and fitness. Physical Functionality Awareness assessed attention to one's physical abilities, and Height Dissatisfaction measured unhappiness with height. Higher scores indicate greater intensity in each aspect.

Data Analysis

All data were analysed using IBM SPSS Statistics Version 20. Descriptive statistics (means, standard deviations, frequencies, percentages) were used to summarise BMI status and body image perceptions. Independent *t*-tests were applied to examine gender differences. Pearson's correlation test was used to assess the association between BMI and body image perception, with statistical significance set at $p < 0.05$.

RESULTS

Table 1 shows the sociodemographic characteristics of the participants. About 140 students participated, of which the majority were female (63.6%), and most were in Form 1 (36.4%) or Form 2 (39.3%). The sample was predominantly Chinese (81.4%), followed by Malay (11.4%) and Indian (5.0%). More than half of the participants (56.4%) reported knowing the ideal BMI value. While only 16.4% reported recent dieting or supplement use for weight loss, most did not engage in such practices. Regarding exercise, 30.0% reported exercising twice a week, and 11.4% reported daily exercise, whereas 7.1% reported never exercising.

Table 1: Descriptive analyses of participants' sociodemographic data

Characteristics	Mean (SD) (%)
Gender	
Male	51 (36.4)
Female	89 (63.6)
Age (year)	
13	51 (36.4)
14	55 (39.3)
15	34 (24.3)
Race	
Malay	16 (11.4)
Chinese	114 (81.4)
Indian	7 (5.0)

Others	3 (2.1)
BMI knowledge	
Know the ideal BMI value	79 (56.4)
Do not know the ideal BMI value	61 (43.6)
Weight loss strategy	
Has recently dieted or taken supplements to lose weight	23 (16.4)
Has not recently dieted nor taken supplements to lose weight	117 (83.6)
Exercise frequency	
Once a week	37 (26.4)
Twice a week	42 (30.0)
More than twice a week	35 (25.0)
Every day	16 (11.4)
Never	10 (7.1)

Table 2 summarises the anthropometric and body composition data. The mean height and weight of participants were 156.6 ± 7.5 cm and 48.5 ± 11.8 kg, respectively, with an average BMI of 19.8 ± 4.2 kg/m². Most students (75.0%) fell within the normal BMI range. However, 10.0% were overweight and 7.1% obese, while 4.3% were classified as thin and 3.6% as severely thin. These findings indicate that while most students maintained a normal BMI, both undernutrition and excess weight coexisted in this population. Female students had a higher mean body fat percentage (23.3 ± 3.8) compared to males (14.9 ± 5.4), while males recorded greater skeletal muscle mass (37.2 ± 5.7 kg vs. 27.6 ± 2.2 kg).

Table 2: Mean of participants' anthropometric data

Anthropometric characteristics	Mean (±SD)
Height (cm)	156.6 (±7.5)
Weight (kg)	48.5 (±11.8)
BMI (kg/m^2)	19.8 (±4.1)
Fat percentage (%)	
Male	14.9 (±5.4)
Female	23.3 (±3.8)
Skeletal muscle mass (kg)	
Male	37.2 (±5.7)
Female	27.6 (±2.2)
BMI classification	**n (%)**
Severe thinness	5 (3.6)
Thinness	6 (4.3)
Normal	105 (75.0)
Overweight	14 (10.0)
Obese	10 (7.1)

Table 3 highlights how participants perceived their body image, presenting mean scores for each of the four domains. Attractiveness Evaluation had the highest mean score (18.98 ± 4.68), suggesting students generally viewed themselves positively in terms of appearance. Negative Affect followed (17.69 ± 6.90), indicating that many still experienced emotional vulnerabilities regarding body image. Physical Functionality Awareness had a moderate mean score (12.56 ± 3.44), reflecting some awareness of physical health and ability, while Height Dissatisfaction was lowest (10.00 ± 3.13), suggesting that participants were generally accepting of their height.

Besides, the mean scores of body image perception domains across different BMI classifications are also shown in Table 3. In the Negative Affect domain, students who were obese or overweight reported the highest levels of dissatisfaction, while those with normal BMI or thinness showed the lowest. This indicates that extreme BMI categories, both high (obese/overweight) and very low (severe thinness), are associated with greater emotional distress regarding body image.

For Attractiveness Evaluation, students in the thinness category reported the highest satisfaction, whereas those in the severe thinness and obese categories recorded the lowest. This suggests that moderate thinness may be perceived positively, but more extreme deviations from normal BMI reduce perceived attractiveness.

In terms of Physical Functionality Awareness, overweight and obese students reported higher awareness of body functionality compared to those with normal or severely thin BMI. This may reflect heightened consciousness about health among students with higher body weight.

Finally, in the Height Dissatisfaction domain, obese and overweight students expressed greater dissatisfaction compared to their normal-weight and severely thin peers, indicating that higher BMI may influence not only weight-related dissatisfaction but also concerns about height.

Table 3: Body image perception according to BMI classification (mean ± SD)

Domain	Total score (mean ± SD)	Severe thinness	Thinness	Normal	Overweight	Obese
Negative Affect	17.69 ± 6.90	19.00 ± 10.20	15.00 ± 4.69	16.45 ± 6.12	21.21 ± 7.12	26.80 ± 5.98)
Attractiveness Evaluation	18.98 ± 4.68	14.80 ± 2.78	21.50 ± 3.15	19.15 ± 4.90	19.14 ± 4.52	17.50 ± 2.17)
Physical Functionality Awareness	12.56 ± 3.44	9.60 ± 3.65	13.33 ± 2.73	12.40 ± 3.46	13.71 ± 3.12	13.60 ± 3.50
Height Dissatisfaction	10.00 ± 3.13	8.80 ± 3.56	10.00 ± 2.97	9.87 ± 3.29	10.50 ± 2.38	11.30 ± 2.11

Based on Table 4, BMI showed a statistically significant moderate positive correlation with Negative Affect (r = 0.338, p < 0.001), indicating that students with higher BMI were more likely to report body dissatisfaction. A weaker but still significant positive correlation was also observed between BMI and Physical Functionality Awareness (r = 0.218, p = 0.010), suggesting that students with higher BMI demonstrated greater awareness of maintaining physical functionality. In contrast, no significant associations were found between BMI and Attractiveness Evaluation (r = 0.015, p = 0.856) or between BMI and Height Dissatisfaction (r = 0.123, p = 0.147).

Table 4: Correlations between BMI and body image perceptions

Characteristics	*r*	*p*-value
BMI and Negative Affect	0.338	0.000*
BMI and Attractiveness Evaluation	0.015	0.856
BMI and Physical Functionality Awareness	0.218	0.010*
BMI and Height Dissatisfaction	0.123	0.147

*p < 0.05, significantly different (2-tailed). Statistical test analysis used: Pearson's Correlation

DISCUSSIONS

Sociodemographic Characteristics of the Participants

Most participants in this study were Chinese (81.4%), followed by Malay (11.4%), Indian (5.0%), and other ethnic groups (2.1%). This ethnic distribution is important, as cultural values and media portrayals influence how adolescents perceive body image. Previous research has shown that Malaysian Chinese adolescents are more likely to internalise thinner body ideals, often leading to dieting and higher dissatisfaction compared to their Malay and Indian peers (Wang et al., 2021; Yeng & Sedek, 2012). Malay adolescents, by contrast, tend to adopt more flexible attitudes shaped by family and cultural expectations (Cheah et al., 2017; Wong & Say, 2013). These findings suggest the need for culturally tailored interventions in Malaysia's diverse adolescent population.

Knowledge of BMI was nearly equally divided between those who reported awareness and those who did not. Although BMI is introduced in school health curricula, comprehension appears limited. Previous studies highlight that Malaysian adolescents often recognise the term "BMI" but lack understanding of its health implications (Musa et al., 2021; UNESCO, 2019). Many rely on peers and social media for health information, which tends to emphasise appearance rather than health outcomes (Muda et al., 2021). This may explain why adolescents misinterpret BMI more as a measure of attractiveness than health.

In addition, some students reported engaging in dieting and supplement use for weight control. This aligns with national data showing widespread use of supplements among Malaysian adolescents (Malaysia School-based Nutrition Survey, 2012). Motivations often stem from parental encouragement, peer influence, or media exposure (Yeo et al., 2014). However, reliance on unregulated weight-loss products raises health concerns, especially as body dissatisfaction and societal pressures drive such practices (Chau et al., 2024).

The majority of students reported engaging in weekly exercise, consistent with NHMS (2017) findings that over half of Malaysian adolescents meet recommended activity levels. Nonetheless, sedentary behaviours remain common, particularly among females, who consistently report lower physical activity participation than males due to sociocultural and environmental barriers (Mohd Salleh et al., 2018). This suggests that interventions promoting physical activity should address both gender-specific barriers and broader lifestyle patterns.

In the present study, most participants were within the normal BMI range, whereas the remaining were classified as underweight or overweight/obese. This reflects the dual burden of malnutrition among Malaysian adolescents, as noted in the NHMS (2022). Despite most students having a normal BMI, the occurrence of overweight and obesity underscores the need for continued attention to adolescent weight management. Excess weight in adolescence is strongly linked to non-communicable diseases and poor mental health later in life (Biswas et al., 2022). The presence of underweight students also highlights that undernutrition persists, underscoring the complexity of adolescent nutrition challenges (Salleh et al., 2024).

Body Image Perceptions According to BSIQ-SF Domains

Overall, adolescents in this study reported moderate satisfaction with their appearance, particularly in the Attractiveness Evaluation domain. However, this was accompanied by higher scores in the Negative Affect domain, indicating emotional distress despite some acceptance of appearance. This duality has been observed in previous studies, where adolescents express both approval and discontent regarding body image, often contributing to poor mental health outcomes (Cash & Smolak, 2011).

Moderate scores in Height Dissatisfaction and Physical Functionality Awareness further suggest that adolescents hold mixed attitudes toward their physical development. Concerns about height, commonly reported during adolescence, often influence self-esteem and peer relationships (Ricciardelli & McCabe, 2004). This finding supports prior research showing that Asian adolescents, including Malaysians, frequently struggle with conflicting feelings about body size and development due to cultural ideals and media pressures (Chen et al., 2019; Liew & Chooi, 2020).

Body Image Perceptions According to BMI Classifications

Body image perceptions varied considerably by BMI classification. Students classified as obese reported the highest levels of Negative Affect, consistent with literature linking higher BMI to body dissatisfaction and psychological distress (Neumark-Sztainer et al., 2006; Puhl & Heuer, 2010). Social stigma and teasing may exacerbate these negative feelings, further impacting well-being (Haines et al., 2006).

By contrast, students in the thin category reported the greatest satisfaction in the Attractiveness Evaluation domain, while those classified as severely thin reported the lowest satisfaction. This aligns with prior findings that, although some underweight adolescents feel content with appearance, others may experience dissatisfaction due to health concerns or social perceptions (Ricciardelli & McCabe, 2001; Smolak, 2011).

In the Physical Functionality Awareness domain, overweight and obese students reported higher scores than those in the normal or severe thinness categories. This may indicate heightened awareness of physical health among adolescents with higher or lower BMI, possibly due to personal experiences or external health advice. Similarly, height dissatisfaction was highest among overweight and obese students, suggesting that concerns about stature may compound weight-related body dissatisfaction. These patterns highlight the layered impact of BMI on body image, where deviations from perceived norms intensify dissatisfaction (Cheah et al., 2017).

Correlation between BMI and Body Image Perception Domains

Correlation analysis confirmed a significant positive association between BMI and Negative Affect, indicating that higher BMI is linked to greater dissatisfaction and negative emotions. This supports prior research that connects BMI with psychological distress and low self-esteem (Neumark-Sztainer et al., 2006; Puhl & Latner, 2007). Experiences of stigma and teasing may reinforce this relationship, perpetuating emotional distress (Haines et al., 2006).

A weaker but significant positive correlation was also observed between BMI and Physical Functionality Awareness. Adolescents with higher BMI may be more

conscious of their physical health, reflecting greater awareness of functionality beyond appearance. Prior studies similarly report that overweight adolescents often recognise the importance of physical fitness, even if participation is limited (Deforche et al., 2015; Lubans et al., 2011).

No significant correlations were found between BMI and Attractiveness Evaluation or Height Dissatisfaction. This suggests that perceptions of attractiveness and height are shaped by cultural and psychological factors beyond weight alone (Grogan, 2016; Nguyen et al., 2020). Thus, while BMI strongly influences negative emotions, it may not directly determine broader aspects of body satisfaction.

However, this study has several limitations. The findings were based on a single school with a predominantly Chinese sample, which may limit generalizability to Malaysia's diverse adolescent population. In addition, the use of self-reported questionnaires may have introduced response bias, particularly due to possible language comprehension issues. Future studies should therefore include larger and more ethnically diverse samples; use validated multilingual tools and explore weight-control behaviours in greater detail to better understand the mechanisms linking BMI and body image.

CONCLUSION

This study demonstrated that BMI status is associated with selected dimensions of body image among high school students, particularly emotional responses toward body appearance and awareness of physical functionality, while no meaningful associations were observed for attractiveness evaluation or height-related perceptions. These findings underscore the relevance of BMI in shaping adolescents' body image experiences and highlight the need for school- and community-level strategies that foster positive body perceptions and encourage healthy lifestyle behaviours. To strengthen the evidence base, future research should incorporate additional determinants such as physical activity, sleep patterns, stress levels, and socioeconomic background, and should involve larger and more diverse samples across multiple schools. Such efforts would enhance the generalisability of findings and better inform targeted interventions aimed at supporting adolescents' overall well-being.

ACKNOWLEDGEMENT

The authors would like to thank students, teachers, the Ministry of Education (MOE), Pahang State Education Department, and the Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia (IIUM) Kuantan, for their cooperation and support. This research was not funded by any grant.

REFERENCES

Avci, D., & Keven Akliman, B. (2018). Gender differences in body image and eating attitudes among adolescents. *Journal of Child and Adolescent Psychiatric Nursing*, *31*(2–3), 65–72. https://doi.org/10.1111/jcap.12213

Biswas, T., Townsend, N., Huda, M. M., Maravilla, J., Begum, T., Pervin, S., ... & Mamun, A. (2022). Prevalence of multiple non-communicable disease risk factors among adolescents in 140 countries: A population-based study. *EClinicalMedicine*, *52*.

Cash, T. F., & Smolak, L. (2011). *Body image: A handbook of science, practice, and prevention* (2nd ed.). Guilford Press.

Chau, K., et al. (2024). Body dissatisfaction and weight loss supplement use among Malaysian adolescents: A cross-sectional study. *Asian Journal of Public Health*, *29*(1), 23–35. https://doi.org/10.5678/ajph.2024.2901

Cheah, W. L., et al. (2017). Cultural influences on body image among Malaysian adolescents: A comparative study. *Asian Journal of Psychiatry*, *27*, 45–52. https://doi.org/10.1016/j.ajp.2017.05.013

Cheah, W. L., Hazmi, H., & Chang, C. T. (2017). Disordered eating and body image issues and their associated factors among adolescents in urban secondary schools in Sarawak,

Chen, H., Jackson, T., & Huang, X. (2019). Body dissatisfaction and psychological distress among adolescents: The role of cultural and media influences in Asia. *Body Image*, *29*, 1–10. https://doi.org/10.1016/j.bodyim.2019.07.005

Deforche, B., De Bourdeaudhuij, I., Tanghe, A., & Hills, A. P. (2015). Physical activity and physical self-concept in adolescents with overweight and obesity: A systematic review. *Obesity Reviews*, *16*(5), 345–355. https://doi.org/10.1111/obr.12249

Farah Wahida, A. W., Mohd Nasir, M. T., & Hazizi, A. S. (2011). Physical activity, eating behaviour and body image perception among young adolescents in Kuantan, Pahang, Malaysia. *Malaysian Journal of Nutrition*, *17*(3), 334–344. [PDF] https://nutriweb.org.my/mjn/publication/17-3/e.pdf

Grogan, S. (2016). *Body image: Understanding body dissatisfaction in men, women and children* (3rd ed.). Routledge.

Haines, J., Neumark-Sztainer, D., Eisenberg, M. E., & Hannan, P. J. (2006). Weight teasing and disordered eating behaviours in adolescents: Longitudinal findings from Project EAT (Eating Among Teens). *Journal of Adolescent Health*, *38*(6), 675–683. https://doi.org/10.1016/j.jadohealth.2005.07.002

Holland, G., & Tiggemann, M. (2016). A systematic review of the impact of the use of social networking sites on body image and disordered eating outcomes. *Body Image*, *17*, 100–110. https://doi.org/10.1016/j.bodyim.2016.02.008

Institute for Public Health (IPH) 2017. (2017). Technical Report National Health and Morbidity Survey (NHMS) 2017: Adolescent Health Survey, Malaysia.

Institute for Public Health (IPH) 2022. (2022). Technical Report National Health and Morbidity Survey (NHMS) 2022: Adolescent Health Survey, Pahang, Malaysia.

Jamani, N. A., Said, A. H., AbdAziz, K. H., & AbdRahman, M. A. (2020). Agreement between body weight perception and body weight status among late adolescents in Kuantan, Malaysia. The Medical Journal of Malaysia, *75*(2), 158–163.

Karastergiou, K., Smith, S. R., Greenberg, A. S., & Fried, S. K. (2012). Sex differences in human adipose tissues – the biology of pear shape. *Biology of Sex Differences*, *3*(1), 13. https://doi.org/10.1186/2042-6410-3-13

Khor, G. L., Zalilah, M. S., Phan, Y. Y., Ang, M., Maznah, B., & Norimah, A. K. (2009). Perceptions of body image among Malaysian male and female adolescents. *Singapore Medical Journal*, *50*(3), 303–311. http://smj.sma.org.sg/5003/5003a11.pdf

Liew, S. M., & Chooi, W. S. (2020). Body image perceptions among Malaysian secondary school students: A mixed methods study. *Malaysian Journal of Public Health Medicine*, *20*(1), 45–53.

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2011). Fundamental movement skills in children and adolescents: Review of associated health benefits. *Sports Medicine*, *40*(12), 1019–1035. https://doi.org/10.2165/11536850-000000000-00000

Malaysia School-based Nutrition Survey. (2012). *Nutrition status and dietary supplement use among Malaysian adolescents*. Ministry of Health Malaysia.

Mohd Salleh, M. N., et al. (2018). Gender differences in physical activity among adolescents in Kuala Lumpur. *Malaysian Journal of Sports Science and Recreation*, *14*(2), 89–98.

Muda, M. S., et al. (2021). Social media influence on adolescent health perceptions in Malaysia. *Journal of Health Communication*, *26*(4), 321–330. https://doi.org/10.1080/10810730.2021.1899876

Musa, R., et al. (2021). Understanding of BMI and health among Malaysian secondary school students. *Health Education Journal*, *80*(5), 567–579. https://doi.org/10.1177/0017896921100421 2

National Academies of Sciences, Engineering, and Medicine. (2019). *The promise of adolescence: Realizing opportunity for all youth*. The National Academies Press. https://doi.org/10.17226/25388

Neumark-Sztainer, D., Paxton, S. J., Hannan, P. J., Haines, J., & Story, M. (2006). Does body satisfaction matter? Five-year longitudinal associations between body satisfaction and health behaviors in adolescent females and males. *Journal of Adolescent Health*, *39*(2), 244–251. https://doi.org/10.1016/j.jadohealth.2005.12.001

Nguyen, T. T., et al. (2020). Cultural perceptions of height and attractiveness in Asian societies. *International Journal of Social Psychology*, *35*(4), 312–324. https://doi.org/10.1080/00207594.2020.1754623

Pon, L. W., Jr, Kandiah, M., & Mohd Nasir, M. T. (2004). Body image perception, dietary practices and physical activity of overweight and normal weight

Malaysian female adolescents. Malaysian journal of nutrition, *10*(2), 131–147.

Puhl, R. M., & Latner, J. D. (2007). Stigma, obesity, and the health of the nation's children. *Psychological Bulletin*, *133*(4), 557–580. https://doi.org/10.1037/0033-2909.133.4.557

Puhl, R. M., & Heuer, C. A. (2010). Obesity stigma: Important considerations for public health. *American Journal of Public Health*, *100*(6), 1019–1028. https://doi.org/10.2105/AJPH.2009.159491

Ricciardelli, L. A., & McCabe, M. P. (2001). Gender differences in body image and eating behaviors: A biopsychosocial perspective. *Journal of Social Psychology*, *141*(4), 493–510. https://doi.org/10.1080/00224540109600569

Ricciardelli, L. A., & McCabe, M. P. (2004). A biopsychosocial model of disordered eating and the

Wang, Y., et al. (2021). Ethnic differences in body image and dieting behaviors among Malaysian adolescents. *Journal of Adolescent Health*, *68*(4), 712–718. https://doi.org/10.1016/j.jadohealth.2020.11.012

Wong, L. P., & Say, K. (2013). Familial and cultural influences on body image among Malay adolescents. *Malaysian Journal of Psychiatry*, *22*(1), 15–22.

World Health Organization. (2007). WHO child growth standards: BMI-for-age (5–19 years). https://www.who.int/tools/growth-reference-data-for-5to19- years/indicators/bmi-for-age.

Yeng, S. S., & Sedek, R. (2012). Body image attitudes among Malaysian adolescents: Ethnic and gender differences. *Asian Journal of Psychiatry*, *5*(4), 300–305. https://doi.org/10.1016/j.ajp.2012.04.007

Yeo, L. S., et al. (2014). Dietary supplement use and body image concerns among Malaysian adolescents. *Asia Pacific Journal of Clinical Nutrition*, *23*(3), 456–462. https://doi.org/10.6133/apjcn.2014.23.3.13

pursuit of muscularity in adolescent boys. *Psychological Bulletin*, *130*(2), 179–205. https://doi.org/10.1037/0033-2909.130.2.179

Salleh, M. N., et al. (2024). Undernutrition among Malaysian adolescents: Prevalence and associated factors. *Malaysian Journal of Public Health Medicine*, *24*(1), 45–53.

Smolak, L. (2011). Body image development in childhood. In T. F. Cash & L. Smolak (Eds.), *Body image: A handbook of science, practice, and prevention* (2nd ed., pp. 67–75). Guilford Press.

Toselli, S.; Zaccagni, L.; Rinaldo, N.; Mauro, M.; Grigoletto, A.; Maietta Latessa, P.; Marini, S. Body Image Perception in High School Students: The Relationship with Gender, Weight Status, and Physical Activity. Children 2023, *10*, 137. https://doi.org/10.3390/ children1001013

UNESCO. (2019). *Health literacy among adolescents in Southeast Asia: A regional report*. UNESCO Bangkok.

ORIGINAL ARTICLE

Understanding Supervisors' Perspective on Clinical Incompetencies and Challenges in Dysphagia Training for Speech-Language Therapy Students

Natrah Ahmad Nordin[1*], Nor Azrita Mohamed Zain[2], Ahmad Aidil Dzulkarnain[3], Norsila Abdul Wahab[4]

[1]Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Pahang, Malaysia
[4]School of Dental Science[1]s, Universiti Sains Malaysia, Kelantan, Malaysia.

Keywords:
Speech-language therapy; dysphagia; training; competency

ABSTRACT

Background: Improvements in clinical education for speech-language therapy students can be made by identifying common areas of clinical incompetency. This study aimed to explore supervisors' perceptions of undergraduate students' clinical incompetencies and challenges during dysphagia training. **Methods:** An online survey was developed and distributed to speech-language therapy supervisors in Malaysia. The survey was designed to explore supervisors' perceptions of students' clinical incompetencies and challenges during dysphagia training. The responses for 4-point Likert scale questions were analysed descriptively. Then, further analysis was done using repeated-measures ANOVA to compare the means of the categories. For open-ended questions, thematic analysis was carried out. **Results:** A total of 11 supervisors participated in the study. Students demonstrated the highest levels of incompetency in the Management and Planning (92.41%) and FEES (83.64%) categories, followed by Case Notes (86.36%) and Case History (83.63%). Despite non-significant statistical differences, effect sizes highlighted FEES and Management and Planning as the most practically important areas of student incompetence.On the other hand, challenges in dysphagia training were categorized into those related to the student and those related to the clinical environment. Contributing factors included limited theoretical understanding, difficulty applying knowledge to practice, low confidence, and restricted clinical exposure. Supervisors recommended enhancing training approaches, strengthening theoretical instruction, and improving access to appropriate facilities and resources. **Conclusion:** The findings highlight key areas where students face difficulties in achieving clinical competence in dysphagia management. Future research should involve larger and more diverse samples, incorporating perspectives from students and academic staff to further inform curriculum development and promote safer, more effective dysphagia management practices.

INTRODUCTION

A speech-language therapist (SLT) is responsible for evaluating, diagnosing, and providing rehabilitation for individuals with speech, language, communication, swallowing, and related disorders (American Speech-Language-Hearing Association [ASHA], 2016). Dysphagia is defined as an impairment of the swallowing process, clinically characterized by abnormal delay or difficulty in moving a food bolus from the oral cavity to the stomach (Clavé & Shaker, 2015). In the management of dysphagia, early and accurate intervention is critical to prevent complications that may lead to serious morbidity or even mortality (Bakhtiyari et al., 2015; Banda et al., 2022; Hazelwood et al., 2023). Consequently, graduates of speech-language therapy programs must demonstrate entry-level competency to ensure patient safety and high-quality care (Attrill et al., 2022). Competency is broadly defined as the integrated capacity of a practitioner to apply knowledge, skills, attitudes, values, and judgment in real clinical environments (Fernandez et al., 2012; McAllister et al., 2011), thereby performing tasks safely, effectively and ethically. In contrast, incompetency refers to the failure to integrate these essential components, resulting in practice that is inconsistent, unsafe, or below the expected professional standard (Overholser & Fine, 1990).

Insufficient clinical skills or inadequate competence in dysphagia management can negatively affect diagnostic accuracy, delay treatment, and compromise clinical

* Corresponding author.
E-mail address: natrahnordin@gmail.com

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

outcomes (Kitila et al., 2024; Vose et al., 2018). These concerns underscore the importance of effective training and supervision during clinical education to support students in developing the competencies required for safe dysphagia practice. Although establishing competency in dysphagia is a fundamental component of speech-language therapy, several studies have reported that SLTs often rate their self-perceived competence in this area as low (Caesar & Kitila, 2020; Mustaffa Kamal, Ward & Cornwell.,2012b; Vose et al., 2018). While much of the existing research has focused on students' or SLTs' self-assessments of knowledge, confidence, and performance in dysphagia management, relatively little attention has been given to the perspectives of clinical supervisors. Supervisors play a crucial role in bridging the gap between classroom learning and clinical application by mentoring, providing feedback, and ensuring that students achieve professional standards of practice (ASHA, 2008). Their perspectives offer valuable input for curriculum development and help identify areas requiring additional support or training. Therefore, this study aims to explore supervisors' perceptions of common clinical incompetencies and the challenges encountered during dysphagia training. The findings are expected to inform improvements in clinical education and contribute to the development of strategies that enhance student preparedness for safe and effective dysphagia management. Fibreoptic endoscopic evaluation of swallowing (FEES) is an instrumental procedure that uses a flexible endoscope to directly visualise pharyngeal swallowing function and detect swallowing abnormalities (Langmore, Kenneth, & Olson, 1988). FEES was included in the questionnaire as it represents the most commonly used instrumental assessment for dysphagia evaluation in Malaysia (Mustaffa Kamal et al., 2012a).

METHODS

Questionnaire

An online survey was developed and validated for this study. The survey consisted of three sections. There were (A) demographic information, (B) Forty 4-point Likert-scale items (ranging from *not at all* to *frequently*) were used to assess possible incompetencies demonstrated by students during clinical training,and (C) open-ended questions of the challenges during training and recommended solutions. Items in section B are divided into seven categories. The categories are (1) case history, (2) oral motor examination (OME), (3) water and food test, (4) FEES, (5) management and plan, (6) case note and (7) professionalism. These items were formulated with reference to past studies and relevant guidelines (Hewat et al., 2020; Macbean et. al., 2013; Kamal Bahri, 2018; Malaysia Ministry of Health, 2023). The survey was validated by three experienced SLTs and one linguist, and was subsequently pretested. The final version of the survey was administered online via Google Forms, and the estimated completion time for participants was approximately 10–15 minutes.

Procedure

The research obtained ethical clearance from the IIUM Research Ethics Committee (IREC) under reference number IREC 2019-151 before commencement. An invitation link containing the research information was distributed via WhatsApp groups of the Speech-Language Pathologists Association of Malaysia (SPEAK), dysphagia workshop group, and speech-language therapy university alumni. The inclusion criteria required respondents to: (1) be practicing SLTs with more than five years of professional experience; (2) have undergone specialized dysphagia training (e.g., certified courses or postgraduate-level training) or have managed dysphagia cases for at least five years; (3) be currently employed in a hospital or university setting; and (4) have experience in supervising speech-language therapy students in dysphagia-related clinical practice. Several SLTs expressed interest in participating; however, responses from individuals who were not directly involved in student supervision were excluded.

Data Analysis

Statistical analyses for the Likert-scale items identifying clinical incompetencies were conducted using the *Statistical Package for the Social Sciences* (SPSS, Version 25). Internal consistency of the items was assessed using Cronbach's alpha, with values ranging from .749 to .883, indicating acceptable to good reliability. Descriptive statistics, including item frequencies and mean scores for each category, were computed. The percentage of high-occurrence incompetencies was calculated by combining responses that fell within the two highest categories of the rating scale. (Dzulkarnain et al., 2019). A repeated-measures analysis of variance (RM-ANOVA) was performed to examine differences in mean error scores across categories, followed by pairwise comparisons with Bonferroni adjustments to identify significant differences between categories. Effect size calculations were also conducted to complement the interpretation of *p*-values. Responses to the open-ended questions were analyzed qualitatively using Braun and Clarke's (2006) thematic analysis approach.

RESULTS

Demographic Information

A total of 11 experienced SLTs met all inclusion criteria and were included in the analysis. Respondents ranged in age from 32 to 42 years (M = 37). Most respondents had more than 10 years of professional experience (n = , 81.9%), and nearly half had over 10 years of supervisory experience (n = 5, 45.5%). All participants were employed in government hospitals or universities, except one who was working in a private institution. On average, SLTs reported spending 18.9 hours (SD = 9.32) per semester on providing dysphagia training and 12.45 hours (SD = 10.06) on fiberoptic endoscopic evaluation of swallowing (FEES) training.

Table 1: Respondent demographic information

Demographic variables		*n*
Education	Bachelor's degree	6
	Masters's degree	5
	PhD	-
Supervising experience	1 to 5 years	1
	5 to 10 years	5
	>10years	5
Institution	University	3
	Government Hospital	7
	Private practice	1
Total of hours spent on dysphagia training (one semester)	<10 hours	3
	11 to 20 hours	3
	>20hours	5
Number of hours spent on FEES training (one semester)	<10hours	6
	11 to 20 hours	1
	>20hours	4

Percentage Occurrence of Clinical Incompetencies Among Speech-Language Therapy Students

The percentage distribution of reported clinical incompetencies among speech-language therapy students is summarized in Table 2. The highest-rated area of difficulty overall was in interpreting oral mechanism examination (OME) findings in relation to cranial nerve function, reported by 90.91% of respondents. Other frequently observed areas of incompetence included difficulty in: 1) deciding if a patient passed or failed the water and food test in relation to OME findings (72.73%); 2) relating FEES findings with cranial nerve functions; 3) integrating the results from all the assessments conducted (72.73%) and; 4) inability to make accurate impression or differential diagnosis (81.89%).

Table 2: Descriptive statistics on types of incompetencies made by students as perceived by supervisors.

No	Item	% (*n*)			
		Not at all	Rarely	Occasionally	Frequently
	CASE HISTORY				
1	Unable to conduct a comprehensive history taking.	-	9.09 (1)	45.45 (5)	45.45 (5)
2	Unable to give simple and clear questions.	9.09 (1)	9.09 (1)	54.55 (6)	27.27 (3)
3	Using jargons and medical terms with the patient and/or caregivers.	-	27.27 (3)	45.45 (5)	27.27 (3)
4	Unable to rephrase questions when patients and/or caregivers find it difficult to understand.	-	27.27 (3)	54.55 (6)	18.18 (2)
5	Unable to probe for details and further data/ facts from patient's given information.	-	-	36.36 (4)	63.64 (7)
	ORAL MOTOR EXAMINATION				
1	Demonstrate wrong technique when examining the patient.	18.18 (2)	18.18 (2)	63.64 (7)	-
2	Unable to give correct/appropriate instructions to the patient.	18.18 (2)	36.36 (4)	36.36 (4)	9.09 (1)
3	Unable to relate the findings with cranial nerve functions.	-	-	9.09 (1)	90.91 (10)
4	Incorrect placement of utensil (e.g. use of tongue depressor).	18.18 (2)	36.36 (4)	45.45 (5)	-
5	Unable to relate the findings to determine the type of instrumental assessment to be conducted.	-	18.18 (2)	36.36 (4)	45.45 (5)
	WATER AND FOOD TEST				
1	Unable to decide which consistency to be prioritized when testing with the patient.	-	9.09 (1)	72.73	18.18 (2)
2	Unable to handle patients with sufficient care to ensure safety.	9.09 (1)	18.18 (2)	18.18 (2)	54.55 (6)
3	Unable to differentiate types of food/drink consistencies.	18.18 (2)	45.45 (5)	27.27 (3)	9.09 (1)
4	Unable to prepare the exact drink/food consistencies using food thickener.	18.18 (2)	18.18 (2)	36.36 (4)	27.27 (3)
5	Unable to decide if a patient passed or failed the test in relation to OME findings.	9.09 (1)	9.09 (1)	9.09 (1)	72.73 (8)
6	Unable to decide indication for instrumental assessment.	-	27.27 (3)	27.27 (3)	45.45 (5)
	FIBERENDOSCOPIC EXAMINATION OF SWALLOWING (FEES)				
1	Unable to give correct/appropriate instructions to the patient.	-	36.36 (4)	9.09 (1)	54.55 (6)
2	Unable to relate the findings with cranial nerve functions.	-	-	27.27 (3)	72.73 (8)
3	Unable to identify the correct anatomical structures.	-	27.27 (3)	9.09 (1)	63.64 (7)
4	Unable to interpret the findings from the observation.	-	9.09 (1)	27.27 (3)	63.64 (7)
5	Unable to give feedback to patient and/or caregivers.	-	9.09 (1)	27.27 (3)	63.64 (7)
	MANAGEMENT AND PLAN				
1	Unable to integrate the results of all the assessments conducted.	-	-	27.27 (3)	72.73 (8)
2	Unable to interpret the findings to determine a correct diagnosis.	-	-	36.36 (4)	63.64 (7)
3	Unable to use the findings to determine a comprehensive plan for the patient.	-	9.09 (1)	45.45 (5)	45.45 (5)

4	Unable to explain the result to the patient.	-	9.09 (1)	63.64 (7)	27.27 (3)
5	Unable to use layman terms when explaining results to the patient.	-	18.08 (2)	36.36 (4)	45.45 (5)
6	Unable to recommend referrals to other professionals if necessary.	-	9.09 (1)	45.45 (5)	45.45 (5)
	CASE NOTE				
1	Unable to come with correct impression or differential diagnosis.	-	9.09 (1)	9.09 (1)	81.89 (9)
2	Unable to write correct notes.	-	18.18 (2)	45.45 (5)	36.36 (4)
3	Unable to write clear notes.	-	18.18 (2)	63.64 (7)	18.18 (2)
4	Unable to write comprehensive notes.	-	9.09 (1)	36.36 (4)	54.55 (6)
	PROFESSIONALISM				
1	Patient was not greeted when entering the room.	36.36 (4)	54.55 (6)	9.09 (1)	-
2	Unable to maintain good manners with the patient.	72.73 (8)	27.27 (3)	-	-
3	Unable to conduct assessment in appropriate sequences.	18.18 (2)	36.36 (4)	27.27 (3)	18.18 (2)
4	Unable to maintain hygiene when dealing with patients.	27.27 (3)	27.27 (3)	27.27 (3)	9.09 (1)
5	Unable to provide consultation using layman terms.	-	18.18 (2)	63.64 (7)	18.18 (2)
6	Unable to maintain a comfortable distance from the patient.	36.36 (4)	54.55 (6)	9.09 (1)	-
7	Unable to appear confident during assessment and result delivery.	-	27.27 (3)	27.27 (3)	45.45 (5)
8	Taking inappropriate time (e.g., too much time or too fast) during assessment.	9.09 (1)	27.27 (3)	36.36 (4)	27.27 (3)

Percentage of high-occurrence incompetencies across categories

The percentage of frequently occurring incompetencies among students was analysed across the various competency categories, with a summary of the findings presented in Figure 1. The *Management and Plan* category demonstrated the highest rate of reported incompetencies (92.41%), followed by the *Case Note* category (86.36%). High levels of incompetency were also noted in the *Case History* (83.63%) and *FEES* (83.64%) domains. Conversely, the *Professionalism* category recorded the lowest frequency of incompetence (39.87%).

Within the Management and Planning category, the highest level of incompetency was related to integrating results from all assessments conducted (72.73%). In the Case Notes category, difficulty in formulating an accurate impression or differential diagnosis was most frequently reported (81.89%). In the FEES category, high levels of incompetency were noted in relating findings to cranial nerve function (72.73%), as well as in identifying anatomical structures, interpreting observational findings, and providing feedback to patients and/or caregivers (each 63.64%).

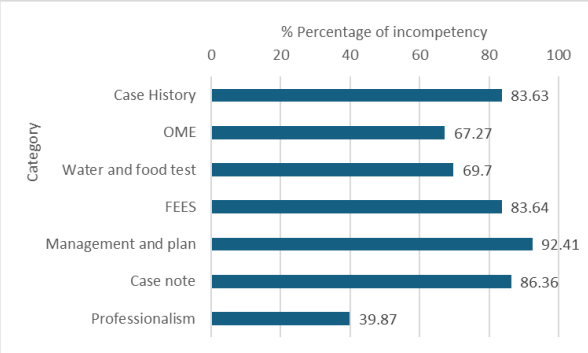


Figure 1: Percentage of high-occurrence incompetencies across categories

Comparison of mean score based on category

To further examine differences in incompetency patterns across categories, mean score comparisons were first conducted (Table 3). The highest mean incompetency scores were observed in the FEES (M = 3.47, SD = 0.63) and Management and Plan (M = 3.42, SD = 0.47) categories, followed by Case Note (M = 3.34, SD = 0.57). In contrast, the lowest mean incompetency score was recorded in the Professionalism category (M = 2.31, SD = 0.53). These descriptive findings were further examined using RM-ANOVA. Although the RM-ANOVA did not reveal statistically significant differences between categories ($p > .005$), the effect sizes from the post-hoc comparisons were predominantly large, indicating meaningful practical differences despite the lack of statistical significance. Notably, large effect sizes were consistently observed in comparisons involving the FEES and Management and Plan categories, reinforcing their status as the most common areas of incompetency.

Table 3: Comparison of mean score based on category

Category	Range	Mean Score	Standard Deviation (SD)
Case history	2.2 – 3.8	3.182	0.510
OME	1.8 – 3.4	2.855	0.580
CA and HLE	2.6 – 4.0	3.273	0.531
Water and food test	1.67 – 4.0	2.985	0.710
FEES	2.2 – 4.0	3.473	0.628
Management and plan	2.5 – 4.0	3.424	0.472
Case note	2.25 – 4.0	3.341	0.573
Professionalism	1.38 – 2.88	2.307	0.530

Table 0: Post-hoc analysis of the eight categories

(I) Category	(J) Category	p-value	Cohen-d	Effect size
Case history	OME	.213	0.59876654	large
	Water and food test	1.00	0.318696803	medium
	FEES	1.0	-	large

		0	0.98068231	
	Management and plan	.493	-0.49250308	medium
	Case note	1.00	-0.29313325	small
	Professionalism	.000*	1.68238 1245	large
OME	Water and food test	1.000	-0.20053468	small
	FEES	.000*	-1.02237202	large
	Management and plan	.001*	-1.07609322	large
	Case note	1.00	-0.84300267	large
	Professionalism	.003*	0.98638717	large
Water and food test	FEES	.110	-0.72808091	large
	Management and plan	.305	-0.72819377	large
	Case note	1.000	-0.55181232	large
	Professionalism	.010*	1.08220 5815	large
FEES	Management and plan	1.000	0.08820 8285	small
	Case note	1.000	0.21958 6682	small
	Professionalism	.000*	2.00664 3942	large
Management and plan	Case note	1.000	0.15811 4887	small
	Professionalism	.000*	2.22581 5146	large
Case note	Professionalism	.007*	-1.87346357	large

Supervisors' Perceptions of Speech-Language Therapy Students' Challenges During Dysphagia Clinical Training

Students' Challenges During Dysphagia Clinical Training

As summarized in Table 5 these challenges were grouped into two main perspectives: (1) challenges related to the students and (2) challenges related to the clinical environment. A majority of supervisors (72.73%, *n* = 8) reported that the primary challenge faced by students was the difficulty in applying theoretical knowledge to clinical practice. One supervisor noted that undergraduates often struggled to "integrate theory into practical aspects and conclude the findings." Additionally, a considerable proportion of respondents identified a lack of theoretical knowledge and low confidence levels (54.55%, *n* = 6, respectively) as recurring issues. For instance, one supervisor observed that students were "not able to explain the normal swallowing process in detail," while another highlighted that undergraduates "lack confidence and are often fearful of handling dysphagic patients, especially those with multiple medical conditions and supportive devices." In regard to challenges related to the clinical environment, most supervisors (45.45%, n = 5) reported that students had limited opportunities to encounter dysphagia patients. One supervisor commented, "unless they are posted in hospitals with many dysphagic patients."

Recommended solutions

Based on the qualitative analysis, three main themes of recommendations emerged, as summarised in Table 6. Qualitative analysis revealed three main themes of recommendations, as summarized in Table 6. Supervisors emphasized increasing clinical training by allocating specific contact hours for dysphagia cases (63.64%, n = 7), placing students in settings with more dysphagia patients (27.27%, n = 3), and enhancing the application of theoretical knowledge to practice (9.09%, n = 1). Strengthening theoretical knowledge was also recommended through independent learning (18.18%, n = 2) and offering adequate dysphagia lectures or classes (18.18%, n = 2), with one supervisor suggesting "two semesters of dysphagia course." Additionally, 18.18% (n = 2) recommended providing advanced learning facilities, including a "dysphagia lab.

Table 5: Challenges faced by undergraduates during dysphagia clinical training

Theme	**Subtheme**	**Examples**	**%** (*n*)
Challenges related to student	Lack of theoretical knowledge	"Not able to explain normal swallowing process in detail."	54.55 (6)
	Lack of application of knowledge into practice	"To integrate theory into practical aspects"	72.73 (8)
	Lack of training in report writing	"Lack of proper training in report writing & documentation"	9.09 (1)
	Lack of confidence	"Lack of confidence and always feel scared of handling dysphagic patient especially those with multiple medical issues…"	54.55 (6)
Challenges related to clinical environment	Lack of opportunity to meet dysphagia patient	"Rarely seen dysphagia case unless they are posting in hospitals with many dysphagic patients."	45.45 (5)
	Lack of opportunity to conduct instrumental assessment	"Lack of hours in observing and conducting instrumental assessments."	18.18 (2)
	Lack of advance tools	"No mannequin provided for students to learn to assess using instruments"	9.09 (1)

Table 6: Recommended solutions

Theme	**Subtheme**	**Examples**	**%** (*n*)
Provide more training	Increase clinical hours for seeing dysphagia patients to graduate	"Ensure that students maintain a logbook with specific hours dedicated to dysphagia, including observation, handling instruments, and managing dysphagia cases."	63.64 (7)
	Attachment at facilities with more dysphagia patients	"To attach with state hospital or national reference hospital."	27.27 (3)
	Enhance application of theoretical knowledge to practice	"University need to teach student how to plan swallowing management from A to Z holistically."	9.09 (1)
Improve theoretical knowledge	Encourage independent learning	"Encourage the use of smartphones to read journal articles and learn about potential causes of dysphagia related to the patient's medical diagnosis."	18.18 (2)
	Provide adequate class/lecture	"Introduce dysphagia management for two semesters."	18.18 (2)
Provide advance facilities/ encourage usage of		"Have a dysphagia lab with mannequin and instruments and tools used in assessment and management of dysphagia."	18.18 (2)

advance facilities

DISCUSSION

This study explored supervisors' perceptions of the common clinical incompetencies, and challenges faced by speech-language therapy students during dysphagia training. The findings revealed that the highest occurrence of incompetencies was in the FEES and management and planning categories.

Difficulties in observation and interpreting FEES are well documented, as the procedure is technically demanding and requires both clinical experience and a strong understanding of anatomical and physiological mechanisms (Colodny, 2002; Langmore et al., 2022; Pilz et al., 2016). The ability to accurately interpret instrumental assessments such as FEES depends on the clinician's capacity to identify relevant anatomical structures and relate the findings to physiological function (Wooi, Scott, & Perry, 2002; Scholten & Russell, 2000). Supervisors in this study also observed that students often lacked sufficient theoretical knowledge, particularly in anatomy and physiology. This observation aligns with previous studies showing that both students and clinicians often struggle to master these foundational concepts (Martin, Bessel, & Scholten, 2014).

Limited clinical exposure to dysphagia cases may also contribute to these challenges, as noted by the supervisors. Opportunities for students to observe or conduct FEES assessments are largely dependent on the availability of equipment and expertise in the clinical setting. This is particularly relevant in Malaysia, where most SLTs primarily manage pediatric language disorders rather than adult dysphagia (Chu et al., 2019; Mustaffa Kamal et al., 2012a). Consequently, students' clinical experience may be influenced by their supervisors' caseloads (McAllister, 2005), potentially leading to variable exposure to dysphagia cases. Resource limitations, inadequate working space, and a shortage of trained educators to conduct instrumental assessments further exacerbate this issue (Mustaffa Kamal et al., 2015).

Frequent incompetencies in the management and planning category suggest that students may struggle to integrate assessment results into accurate impressions and individualized care plans, reflecting gaps in clinical reasoning and the application of theoretical knowledge (Brackenbury et al., 2014; Scholten, 2001). When theoretical and practical training are not well integrated, students may find it challenging to translate classroom learning into clinical practice (Brackenbury et al., 2014). These findings highlight the need for more practice-based training, consistent with Kolb's (1984) experiential learning theory, which emphasizes learning through experience.

The importance of practical learning in speech-language therapy training has been highlighted, with innovations in clinical placements and supervisory models described or proposed to support competency development (Joginder Singh et al., 2019; Sheepway et al., 2011). Programs in Malaysia and internationally have begun to incorporate active learning components within coursework (Ahmad Nordin, 2025; Affoo et al., 2020; Burda & Hageman, 2015; Burda et al., 2022). However, consideration remains as to whether current dysphagia-related coursework sufficiently supports students in managing the complexity of clinical practice. Vose et al. (2018) reported that many universities offer only a single swallowing course, which may be associated with more limited opportunities for in-depth learning. Consistent with this, supervisors in the current study recommended expanding coursework to better support the development of dysphagia-related competencies.

To address limitations in exposure to dysphagia patients and instruments, supervisors emphasized the need for a specific logbook to document encounters, as well as the provision of advanced tools and facilities to enhance hands-on experience. Research suggests that extended dysphagia training and the use of a dysphagia lab can improve students' confidence and competency (Graville, Palmer, & Gorsek, 2015; Wagner & Malandraki, 2016). However, access to such resources may be limited in our context, and these approaches can be resource-intensive. Therefore, more sustainable alternatives, such as simulation studies, may provide additional opportunities for students to gain practical experience in a controlled and safe environment. ,

CONCLUSIONS AND LIMITATIONS

This study identified key areas of clinical incompetency and challenges faced by undergraduate speech-language therapy students during dysphagia training. Supervisors reported the highest incompetency rates in FEES interpretation and management planning, reflecting gaps in applying theoretical knowledge to clinical reasoning. Limited theoretical understanding, low confidence, and restricted clinical exposure were among the main contributing factors.

The findings highlight the need to enhance dysphagia education through additional coursework and improved access to dysphagia cases and instrumental assessment facilities to strengthen practical competence. However, the small sample size and reliance on self-reported data may limit generalizability. Future studies should include larger and more diverse samples, as well as student and academic perspectives, to broaden understanding. Overall, these findings offer important insights to guide

curriculum enhancement and promote safer, more effective dysphagia management practice.

ACKNOWLEDGEMENT

This research was not funded by any grant. The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

Adams, S. N., & Coutts, K. (2025). The use of simulation in the assessment and management of dysphagia: A scoping review. *Clinical Simulation in Nursing, 98*, Article 101661. https://doi.org/10.1016/j.ecns.2024.101661

Ahmad Nordin, N. (2025). *The Development of An Interactive Self-training System for Dysphagia Management* [Unpublished doctoral dissertation]. International Islamic University Malaysia.

Affoo, R., Bruner, J., Dietsch, A., Nellenbach, C., Jones, T., & Lehman, M. (2020). The Impact of Active Learning in a Speech-Language Pathology Swallowing and Dysphagia Course. *Teaching and Learning in Communication Sciences & Disorders*, *4*(2). https://doi.org/10.30707/TLCSD4.2/POPG6689

American Speech-Language-Hearing Association. (2008). *Clinical supervision in speech-language pathology* [Technical report]. https://www.asha.org/policy

American Speech-Language-Hearing Association. (2016). *ASHA's interprofessional education/interprofessional practice (IPE/IPP) activities and collaborations.* Retrieved from http://www.asha.org/Practice/IPE-IPP-Activitiesand-Collaboration

Attrill, S. L., McAllister, S., & Brebner, C. (2022). Not too little, not too much: supervisor perceptions of work-readiness of speech-language pathology graduates. *Advances in Health Sciences Education*, *27*(1), 87–106. https://doi.org/10.1007/s10459-021-10073-8

Bakhtiyari, J., Sarraf, P., Nakhostin-Ansari, N., Tafakhori, A., Logemann, J., Faghihzadeh, S., & Harirchian, M. H. (2015). Effects of early intervention of swallowing therapy on recovery from dysphagia following stroke. In Iran J Neurol (Vol. 14, Issue 3). http://ijnl.tums.ac.ir

Banda, K. J., Chu, H., Kang, X. L., Liu, D., Pien, L. C., Jen, H. J., Hsiao, S. T. S., & Chou, K. R. (2022). Prevalence of dysphagia and risk of pneumonia and mortality in acute stroke patients: A meta-analysis. *BMC Geriatrics, 22*(1), Article 60. https://doi.org/10.1186/s12877-022-02960-5

Brackenbury, T., Folkins, J. W., & Ginsberg, S. M. (2014). Examining educational challenges in communication sciences and disorders from the perspectives of signature pedagogy and reflective practice. *Contemporary Issues in Communication Science and Disorders, 41*, 70–82.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Burda, A., Brotherthon, J., & Duitscher, K. (2022). The cranial nerve exam: Effectiveness of peer-to-peer teaching and experiential learning in speech-language pathology graduate students. *Teaching and Learning in Communication Sciences & Disorders, 6*(3). https://doi.org/10.30707/tlcsd6.3.1664996985.093189

Burda, A., & Hageman, C. (2015). Problem-based learning in speech-language pathology: Format and feedback. *Contemporary Issues in Communication Science and Disorders, 42*, 47–71. https://doi.org/10.1044/cicsd_42_S_47

Caesar, L. G., & Kitila, M. (2020). Speech-Language Pathologists' Perceptions of Their Preparation and Confidence for Providing Dysphagia Services. *Perspectives of the ASHA Special Interest Groups*, *5*(6), 1666–1682. https://doi.org/10.1044/2020_persp-20-00115

Chu, S. Y., Khoong, E. S. Q., Ismail, F. N. M., Altaher, A. M., & Razak, R. A. (2019). Speech-language pathology in Malaysia: Perspectives and challenges. *Perspectives of the ASHA Special Interest Groups, 4*(5), 1162–1166. https://doi.org/10.1044/2019_pers-sig17-2019-0005

Clavé, P., Shaker, R. Dysphagia: current reality and scope of the problem. *Nat Rev Gastroenterol Hepatol* **12**, 259–270 (2015). https://doi.org/10.1038/nrgastro.2015.49

Colodny, N. (2002). Interjudge and intrajudge reliabilities in fiberoptic endoscopic evaluation of swallowing (FEES®) using the penetration–aspiration scale. *Dysphagia, 17*(4), 308–315. https://doi.org/10.1007/s00455-002-0073-0

Dzulkarnain, A. A. A., Rahmat, S., Ismail, A. W., Musa, R., Badzis, M., & Tengku Zam Zam, T. Z. A. (2019). A novel computer-based simulated learning environment in audiology with learning assistance: Preliminary findings. *Medical Journal Malaysia*, *74*(2), 168–173.

Fernandez, N., Dory, V., Ste-Marie, L. G., Chaput, M., Charlin, B., & Boucher, A. (2012). Varying conceptions of

competence: An analysis of how health sciences educators define competence. In *Medical Education* (Vol. 46, Issue 4, pp. 357–365). https://doi.org/10.1111/j.1365-2923.2011.04183.x

Fuller, F. M. (2021). *Speech-language pathologist confidence in dysphagia management* (Unpublished master's thesis). Arkansas State University. https://arch.astate.edu/all-etd/330

Graville, D. J., Palmer, A. D., & Gorsek, S. (March, 2015). Teaching dysphagia skills to graduate students: Experiences with a competency-based program. Poster presentation at the annual meeting of the Dysphagia Research Society, Chicago, IL.

Hazelwood, R. J., Armeson, K. E., Hill, E. G., Bonilha, H. S., & Martin-Harris, B. (2023). Relating Physiologic Swallowing Impairment, Functional Swallowing Ability, and Swallow-Specific Quality of Life. *Dysphagia*, *38*(4), 1106–1116. https://doi.org/10.1007/s00455-022-10532-3

Hewat, S., Penman, A., Davidson, B., Baldac, S., Howells, S., Walters, J., Purcell, A., Cardell, E., McCabe, P., Caird, E., Ward, E., & Hill, A. E. (2020). A framework to support the development of quality simulation-based learning programmes in speech–language pathology. *International Journal of Language and Communication Disorders*, *55*(2), 287–300. https://doi.org/10.1111/1460-6984.12515

Joginder Singh, S., Tan, T. M., & Mustaffa Kamal, R. (2019). Speech-language pathology clinical education in Malaysia: Current practice and future direction. Speech, Language and Hearing. https://doi.org/10.1080/2050571X.2019.1585680

Kamal Bahri, N. (2018). *Development and validation of questionnaire on audiology students' clinical errors* (Unpublished bachelor's dissertation). International Islamic University Malaysia.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.

Kitila, M., Borders, J. C., Krisciunas, G. P., McNally, E., & Pisegna, J. M. (2024). Confidence, Accuracy, and Reliability of Penetration-Aspiration Scale Ratings on Flexible Endoscopic Evaluations of Swallowing by Speech Pathologists. *Dysphagia*, *39*(3), 504–513. https://doi.org/10.1007/s00455-023-10635-5

Kwan, C. C. Y., Weir, K. A., Fong, R., Pereira, V. J., & Law, T. (2025). *Integration of team-based learning and simulation-based learning in clinical education of infant feeding and swallowing assessment and management of speech-language pathology students: A retrospective pre-post intervention study.* BMC Medical Education, 25, 1228. https://doi.org/10.1186/s12909-025-07721-7

Langmore, S.E., Kenneth, S.M.A. & Olsen, N. Fiberoptic endoscopic examination of swallowing safety: A new procedure. *Dysphagia* **2**, 216–219 (1988). https://doi.org/10.1007/BF02414429

Langmore, S. E., Scarborough, D. R., Kelchner, L. N., Swigert, N. B., Murray, J., Reece, S., Cavanagh, T., Harrigan, L. C., Scheel, R., Gosa, M. M., & Rule, D. K. (2022). Tutorial on clinical practice for use of the fiberoptic endoscopic evaluation of swallowing procedure with adult populations: Part 1. *American Journal of Speech-Language Pathology, 31*(1), 163–187. https://doi.org/10.1044/2021_AJSLP-20-00348

Macbean, N., Theodoros, D., Davidson, B., & Hill, A. E. (2013). Simulated learning environments in speech-language pathology: An Australian response. *International Journal of Speech-Language Pathology, 15*(3), 345–357.https://doi.org/10.3109/17549507.2013.779024

Malaysia Ministry of Health. (2023). *Prosedur operasi standard penjagaan pesakit dengan kecelaruan pengambilan makanan (feeding) dan kecelaruan penelanan.* https://hq.moh.gov.my/alliedhealth/index.php/ms/penerbitan/prosedur-operasi-standard

McAllister, S., Lincoln, M., Ferguson, A., & McAllister, L. (2011). A systematic program of research regarding the assessment of speech-language pathology competencies. *International Journal of Speech-Language Pathology*, *13*(6), 469–479. https://doi.org/10.3109/17549507.2011.580782

Mustaffa Kamal, R., Ward, E. C., Cornwell, P., & Sharma, S. (2015). Provision of dysphagia services in a developing nation: Infrastructural challenges. *International Journal of Speech-Language Pathology, 17*(6), 594–604. https://doi.org/10.3109/17549507.2015.1026276

Mustaffa Kamal, R., Ward, E., & Cornwell, P. (2012a). Dysphagia management practices among speech-language pathologists in Malaysia. *Journal of Speech, Language and Hearing Research, 15*(2), 111–128. https://doi.org/10.1179/jslh.2012.15.2.111

Mustaffa Kamal, R., Ward, E., & Cornwell, P. (2012b). Dysphagia training for speech-language pathologists: Implications for clinical practice. *International Journal of Speech-Language Pathology, 14*(6), 569–576. https://doi.org/10.3109/17549507.2012.713394

Overholser, J. C., & Fine, M. A. (1990). Defining the boundaries of professional competence: Managing subtle cases of clinical incompetence. Professional Psychology: Research and Practice, 21(6), 462-469.

Pilz, W., Vanbelle, S., Kremer, B., van Hooren, M. R., van Becelaere, T., Roodenburg, N., & Baijens, L. W. J. (2016). Observers' agreement on measurements in fiberoptic endoscopic evaluation of swallowing. *Dysphagia, 31*(2), 180–187. https://doi.org/10.1007/s00455-015-9673-7

Scholten, I. (2001). Teachers' conceptions of their role in improving students' preparation for clinical work in dysphagia. *American Journal of Speech-Language Pathology, 10*(4), 343–357. https://doi.org/10.1044/1058-0360(2001/029)

Scholten, I., & Russell, A. (2000). Learning about the dynamic swallowing process using an interactive multimedia program. *Dysphagia, 15*(1), 10–16. https://doi.org/10.1007/s004559910004

Vose, A. K., Kesneck, S., Sunday, K., Plowman, E., & Humbert, I. (2018). A survey of clinician decision making when identifying swallowing impairments and determining treatment. *Journal of Speech, Language, and Hearing Research*, *61*(11), 2735–2756. https://doi.org/10.1044/2018_JSLHR-S-17-0212

Wagner, E. M., & Malandraki, G. A. (2016). The development and impact of a university-based dysphagia research clinic in the education of graduate students. Perspectives of the ASHA Special Interest Groups, 1(13), 113–121. https://doi.org/10.1044/persp1.SIG13.113

Wooi, M., Scott, A., & Perry, A. (2001). Teaching speech pathology students the interpretation of videofluoroscopic swallowing studies. *Dysphagia, 16*(1), 32–39. https://doi.org/10.1007/s00455000004

ORIGINAL ARTICLE

Cognitive Interference Effects of Colour-Word and Counting Stroop Task on the Auditory Brainstem Response (ABR) in Children and Adults

Nur Hafizah Sulaiman[1,2], Muhamad Kamarul Asyraf Bakri[1], Muhammad Imran Yusri[1], Rina Kartini Alhawari[1], Ahmad Aidil Arafat Dzulkarnain[1,2*]

[1]Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia
[2]Children Health and Wellbeing Research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

Keywords:
Sensory gating; auditory sensory gating; auditory brainstem response; colour-word Stroop task; counting Stroop task

ABSTRACT

Background: Auditory sensory gating refers to the brain's ability to suppress irrelevant auditory information from further processing. Recent studies have explored the use of the counting Stroop task, performed concurrently with the auditory brainstem response (ABR) test, to assess sensory gating at the brainstem level. However, the classical colour-word Stroop task, which is known to elicit stronger cognitive interference, has not yet been examined in conjunction with ABR testing. Therefore, this study aimed to compare the effects of cognitive interference induced by the colour-word and counting Stroop tasks on ABR outcomes. **Methods:** Seven children and eleven adults with normal hearing participated in this study. Each participant underwent ABR testing concurrently with both Stroop task types (colour-word and counting). For each task, two Stroop conditions (incongruent and neutral) were administered. Four measurement parameters (Stroop interference of percentage of correct responses and response time, and cognitive interference of ABR wave V amplitude and latency) were analysed using within-group comparisons to examine differences between the two types of Stroop task. **Results:** No statistically significant differences were observed in any measurement parameters between the two types of Stroop task for either age group. However, medium to large effect sizes were found among children in Stroop interference of response time and cognitive interference of wave V amplitude. **Conclusion:** Although no significant differences were identified, the colour-word Stroop task appeared to elicit slightly stronger cognitive interference in children, potentially reflecting greater inhibitory control demands. Clinically, these findings suggest that combining the colour-word Stroop task with ABR testing may provide a more sensitive approach for assessing inhibitory control of sensory gating at the brainstem level in children and for identifying those with atypical inhibitory control abilities.

INTRODUCTION

Sensory gating is a fundamental neural mechanism, which regulates the brain's sensitivity to incoming stimuli by selectively filtering out redundant and irrelevant information to prevent sensory overload, thereby allowing attentional and cognitive resources to be efficiently allocated to relevant stimuli (Jones et al., 2016). Although sensory gating occurs across multiple sensory modalities, in the auditory domain, it is referred to as auditory sensory gating. This term reflects the brain's ability to filter out repetitive or unwanted auditory stimuli to maintain optimal auditory processing efficiency (Dzulkarnain et al., 2021).

Auditory sensory gating can be evaluated using various assessment methods, including perceptual rating questionnaires, behavioural psychological assessment, and electrophysiological recordings of auditory evoked potential (AEP) (Dzulkarnain et al., 2021,). Among these, the most established method involves the paired-click paradigm during electroencephalography (EEG), which assesses the suppression of event-related potentials (ERPs), such as P50, N100, and P200. This method primarily measures cortical sensory gating, involving cortical brain regions such as the superior temporal gyrus (STG) and Heschl's gyrus (Knott et al., 2009; Korzyukov et al., 2007; Potter et al., 2006; Villarreal & Hunter, 2015).

To assess sensory gating at the subcortical level, particularly within the auditory brainstem, recent research has combined the auditory brainstem response (ABR) test with concurrent cognitive tasks such as the Stroop task (Brännström et al., 2018; Dzulkarnain et al., 2020, 2021;

* *Corresponding author.*
E-mail address: ahmadaidil@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

Shahrudin et al., 2023). These studies have primarily employed the counting Stroop task, in which participants view digits on a screen, count the number of digits presented, and press the corresponding number key on the keyboard. During the incongruent condition of the counting Stroop task, where the number of digits displayed does not match its numerical value (e.g., "11", "222", "3333", "4", etc.), cognitive interference arises due to conflict between the automatic process of naming the digit and the task-required process of counting the digits. This cognitive interference increases cognitive load and engages inhibitory control mechanisms associated with sensory gating (Hershman et al., 2024; Scarpina & Tagini, 2017). Consequently, activation of the sensory gating system suppresses the processing of irrelevant stimuli, including the auditory stimuli presented during the concurrently recorded ABR, leading to reduced neural responsiveness as reflected by a decrease in ABR wave V amplitude (Brännström et al., 2018; Dzulkarnain et al., 2021; Shahrudin et al., 2023; Sörqvist et al., 2012). This cross-modal suppression, involving inhibitory interactions between visual and auditory modalities, is referred to as cross-sensory gating (Magnée et al., 2009; Schramm et al., 2023).

While the counting Stroop task has been frequently employed in previous ABR-based studies to evaluate sensory gating at the brainstem level, the colour-word Stroop task, which is the classical and most commonly used variant, has not yet been examined in this context. Recent meta-analytic evidence by Müller et al. (2024) indicates that the colour-word Stroop task induces the strongest cognitive interference effects on behavioural performance, as reflected in longer response time compared to other Stroop variants. Given this finding, it remains unclear whether the stronger cognitive interference elicited by the colour-word Stroop task would lead to greater modulation of subcortical auditory responses as measured by the ABR.

Therefore, the present study aimed to compare the effects of cognitive interference induced by the colour-word Stroop task and the counting Stroop task on ABR outcomes in two age groups: children and adults. By exploring both behavioural and electrophysiological indices, this study sought to determine whether differences in task-related cognitive interference are reflected not only in behavioural performance measures but also in neural responses at the auditory brainstem level.

METHODOLOGY

Ethics

Ethical approval for this study was obtained from the Research Ethics Committee of the International Islamic University Malaysia (IIUM) (Ref: KAHS 101/24). Prior to participation, written informed consent was secured, following a full disclosure of the study's objectives, procedures, potential risks and benefits, and anticipated outcomes.

Participants

Seven children (four males, three females), aged 10 to 17 years, and eleven adults (six males, five females) with normal hearing participated in this study. Children aged 10 years old and older were recruited because sensory gating is expected to mature by approximately 8 years of age, and the reading fluency required to read the color word in the color-word Stroop task is typically fully achieved by around 9 years of age.

Procedure

Preliminary hearing assessments

To ensure normal hearing sensitivity, all participants completed hearing assessment, which was conducted using an Interacoustics AC40 pure-tone audiometer to measure audiometric thresholds and a Grason-Stadler TympStar Pro middle ear analyser to evaluate middle-ear compliance. Participants were classified as having normal hearing if both ears met all of the following criteria: (i) pure-tone audiometric thresholds of 20 dB HL or better from 250 to 8000 Hz, (ii) a type A tympanogram with peak compliance between 0.25 to 1.05 (children) or 0.3 to 1.7 (adults) ml, peak pressure between -100 to +50 (children) or ±50 (adult) decapascal (daPa), and ear canal volume between 0.3 to 0.9 (children) or 0.9 to 2.0 (adult) cm^3, and (iii) the presence of acoustic reflex at 1000 Hz.

Development of counting and colour-word Stroop task

The counting and colour-word Stroop tasks were developed using E-prime software to present the Stroop trials, record responses, and analyse participants' performance. Each Stroop task consisted of two conditions: incongruent and neutral.

For the counting Stroop task, digits (1 – 4) were used as test stimuli. In the incongruent counting condition, digits were displayed in quantities that did not match their numerical value (e.g., "11", "111", "1111", "2", "222", "2222", "3", "33", "3333", "4", "44", and "4444").

Participants were instructed to count the number of digits displayed, rather than identify the digit itself, and to enter their responses using a wireless numerical keypad to minimise movement-related artifacts during ABR recording. Only the 1 to 4 keys were programmed as response options in E-prime, while other keys were deactivated. For example, the "1" key corresponded to a single digit (e.g.: "1"), the "2" key to two digits (e.g.: "11"), the "3" key to three digits (e.g.: "111") and the "4" key to four digits (e.g.: "1111"). In the neutral counting condition, the symbol "#" was displayed instead of digits, appearing one to four times (e.g., "#", "##", "###", and "####"). Participants were instructed to count the number of symbols and respond using the same keypad configuration.

For the colour-word Stroop task, colour names in Malay ("MERAH", "BIRU", "HIJAU", and "KUNING") were used as stimuli. In the incongruent colour-word condition, each word was presented in a colour different from its meaning (e.g., "**MERAH**" shown in blue, "**BIRU**" shown in green, etc.). Participants were instructed to identify the colour of the word rather than read it and respond via the same numerical keypad. The keypad was programmed as follows: "1" for blue, "2" for green, "3" for red, and "4" for yellow. To minimise additional cognitive load from the arbitrary association between the numeric keys and the colours, coloured-stickers corresponding to each colour were placed on the key. In the neutral colour-word condition, the stimulus "XXXX" was presented in the same four colours (red, blue, green, and yellow), serving as the baseline measurement. Participants were required to identify the colour of the "XXXX" stimulus using the keypad.

Each condition of each Stroop task comprised 60 randomised trials. Stimuli were presented at the centre of the screen for a maximum of 4000 ms, preceded by a 1000-ms fixation cross. The next item appeared automatically after a response or once the time limit was reached.

ABR test parameters

ABRs were recorded using an Interacoustics Eclipse system. Narrowband (NB) CE-chirps centred at 1 kHz were presented at 70 dBnHL via insert earphones. Stimuli were delivered at a rate of 27.7 Hz with alternating polarity. Responses were recorded over a 20-ms time window using both ipsilateral and contralateral montages. Electrodes were placed with the non-inverting electrode on the high forehead, inverting electrodes on both earlobes, and the ground electrode on the nape of the neck. Contralateral recording was employed to facilitate identification of wave V, as ipsilateral montages often result in fused wave IV and V complexes (Dzulkarnain et al., 2014). Bayesian weighted averaging was applied, and artifact rejection was set at ±40 µV to minimise noise contamination. Filters were set at 30 Hz (high-pass) and 3000 Hz (low-pass) to capture the spectral energy of the ABR, which typically ranges between 100 and 1000 Hz (British Society of Audiology [BSA], 2019).

The number of averages was not fixed and varied, depending on the duration each participant took to complete the Stroop task. ABR recording continued until the participant had responded to all task items, resulting in variable recording durations and consequently variable numbers of averages. Nevertheless, a minimum of 2000 sweeps was ensured for each ABR recording, in line with conventional signal-averaging requirements (BSA, 2019). To ensure reliability and comparability, only recordings with at least 2000 averages and a multiple points F-statistics (FMP) ratio ≥ of 3:1 was included in the analysis.

ABR with Stroop task

Prior to ABR testing with the Stroop task, participants completed a 20-item practice trial for each Stroop task, containing an equal mix of incongruent and neutral trials. Written feedback was displayed on the screen after each trial, and participants were required to achieve at least 80% accuracy before proceeding. If this criterion was not met, the practice trial automatically restarted.

ABRs were then recorded during each of the Stroop conditions (incongruent and neutral) for both types of Stroop task (counting and colour-word). For each condition, two ABR recordings were obtained to verify waveform repeatability, resulting in eight recordings per participant. All recordings were conducted while participants remained seated and focused on the screen. The order of the Stroop task type and conditions was randomised across participants.

Following data collection, two behavioural outcomes (response time and percentage of correct responses) were retrieved from E-prime's E-Merge and E-DataAid applications. The E-prime software automatically recorded the response time for each trial, which was then averaged for each condition of the Stroop task. Participant responses were automatically logged and later manually verified against the correct answers to calculate a percentage score of correct responses. Using the average response time and percentage of correct responses from the incongruent and neutral conditions, Stroop interference of response time and percentage of correct responses was calculated for each type of Stroop task by subtracting the neutral condition's value from the incongruent condition's value (see Eq. (1) and Eq. (2)). The

Stroop interference is typically used as a measurement of inhibitory control ability (Forte et al., 2024).

Stroop interference of response time (1)

Stroop interference of percentage of correct responses (2)

For ABR outcomes, waveforms were visually inspected by two experienced clinical audiologists. Latencies and amplitudes of wave V were retrieved, and cognitive interference of wave V amplitude and latency was calculated for each type of Stroop task by subtracting the neutral condition values from those obtained in the incongruent condition (see Eq. (3) and Eq. (4).

Cognitive interference of wave V amplitude (3)

Cognitive interference of wave V latency (4)

Data Analysis

Data were first examined for normality prior to inferential analysis. Normality was evaluated through visual inspection of histograms, evaluation of skewness (acceptable range: ±1) and kurtosis (acceptable range: ±3), and the Shapiro-Wilk test ($p > 0.05$) as described by Sharif (2019). Based on normality results, either parametric or non-parametric tests were applied in the subsequent inferential analyses. Within-group comparisons of Stroop interference (percentage of correct responses and response time) and cognitive interference (ABR wave V amplitude and latency) between colour-word and counting Stroop tasks were analysed using either a paired-sample t-test or a Wilcoxon signed-rank, as appropriate. Effect sizes were calculated using Cohen's *d* for the paired-sample t-test and interpreted as small (0.2), medium (0.5), or large (0.8), or using the *r* for the Wilcoxon signed-rank test and interpreted as small (0.1 - < 0.3), medium (0.3 - < 0.5), or large (≥ 0.5).

RESULTS

Differences in Stroop Interference and Cognitive Interference of the ABR Test in Adults

Table 1 presents the within-group comparisons of Stroop interference of percentage of correct responses and response time, as well as cognitive interference of wave V amplitude and latency, between the colour-word and counting Stroop tasks in adults.

A Wilcoxon signed-rank test was conducted to compare the Stroop interference of the percentage of correct responses between the colour-word Stroop task (*Mdn* = 0.99, *IQR* = 1.48) and the counting Stroop task (*Mdn* = 0.98, *IQR* = 1.53). The results indicated no significant difference between the two tasks, $Z = -0.965$, $p = 0.287$, $r = 0.29$, consistent with a medium effect size. A paired sample t-test was conducted to compare the Stroop interference of response time between using the colour-word Stroop task ($M = 97.82$, $SD = 10.59$) and the counting Stroop task ($M = 55.83$, $SD = 61.28$). No significant difference was also found, $t(10) = 0.977$, $p = 0.352$, $d = 0.30$, consistent with a small to medium effect size.

For ABR cognitive interference, Wilcoxon signed-rank tests were conducted to compare both measurement parameters (wave V amplitude and latency) between the colour-word and counting Stroop tasks. The results indicated no significant differences in either measurement parameter between the two tasks (cognitive interference of wave V amplitude: colour-word, *Mdn* = -0.015, *IQR* = 0.098; counting, *Mdn* = -0.009, *IQR* = 0.064; $Z = -0.267$, $p = 0.790$, $r = 0.08$; cognitive interference of wave V latency: colour-word, *Mdn* = -0.06, *IQR* = 0.123; counting, *Mdn* = -0.01, *IQR* = 0.105; $Z = -0.578$, $p = 0.563$, $r = 0.17$), consistent with small effect sizes for both measures.

Differences in Stroop Interference and Cognitive Interference of the ABR Test in Children

Table 2 presents the within-group comparisons of Stroop interference of percentage of correct responses and response time, as well as cognitive interference of wave V amplitude and latency, between the colour-word and counting Stroop tasks in children.

A paired-sample t-test was conducted to compare Stroop interference for both measures (percentage of correct responses and response time) between the colour-word and counting Stroop tasks. The results indicated no significant differences in either measurement parameter between the two tasks (Stroop interference of percentage of correct response: colour-word, $M = 0.99$, $SD = 0.08$

counting, *M* = 0.98, *SD* = 0.040; *t*(6) = 0.81, *p* = 0.45, *d* = 0.31; Stroop interference of response time: colour-word, *M* = 163.51, *SD* = 240.35; counting, *M* = 93.09, *SD* = 112.65; *t*(6) = 2.23, *p* = 0.059, *d* = 0.84), indicating a small to medium effect size for percentage of correct response and a large effect size for response time.

For the ABR cognitive interference, a Wilcoxon signed-rank test was conducted to compare the cognitive interference of wave V amplitude between the colour-word (*Mdn* = -0.02, *IQR* = 0.42) and counting Stroop tasks (*Mdn* = 0.38, *IQR* = 0.39). The results indicated no significant difference between the two tasks, *Z* = -1.104, *p* = 0.32, *r* = 0.42, although a medium effect size was observed. A paired-samples t-test was conducted to compare the cognitive interference of wave V latency between the colour-word (*M* = 0.154, *SD* = 0.244) and the counting Stroop Task (*M* = 0.016, *SD* = 0.184). The results indicated no significant difference between the two tasks, *t*(6) = 1.10, *p* = 0.32, *d* = 0.42, consistent with a small to medium effect size.

Table 1: Within-group comparisons of Stroop interference and cognitive interference between the colour-word and counting Stroop task in adults

Interference Index	Types of Stroop Task	Mean (SD)* / Median (IQR)**	Test Statistics	*p*-value	Effect Size (*d* / *r*)	Interpretation
Stroop interference of percentage of correct responses	Colour-word	0.99 (1.48)**	*Z*(10) = –0.965	0.287	*r* = 0.29	Small
	Counting	0.98 (1.53)**				
Stroop interference of response time	Colour-word	97.82 (103.59)*	*t*(10) = 0.977	0.352	*d* = 0.30	Small–medium
	Counting	55.83 (61.28)*				
Cognitive interference of wave V amplitude	Colour-word	–0.015 (0.098)**	*Z*(10) = –0.267	0.790	*r* = 0.08	Small
	Counting	–0.009 (0.064)**				
Cognitive interference of wave V latency	Colour-word	–0.06 (0.123)**	*Z*(10) = –0.578	0.563	*r* = 0.17	Small
	Counting	–0.01 (0.105)**				

Note: * = value for Mean (SD); ** = value for Median (IQR); SD = standard deviation; IQR = interquartile range

Table 2: Within-group comparisons of Stroop interference and cognitive interference between the colour-word and counting Stroop task in children

Interference Index	Types of Stroop Task	Mean (SD)* / Median (IQR)**	Test Statistics	*p*-value	Effect Size (*d* / *r*)	Interpretation
Stroop interference of percentage of correct responses	Colour-word	0.039 (0.089)*	*t*(6) = 0.81	0.45	0.31	Small–medium
	Counting	-0.033 (0.040)*				
Stroop interference of response time	Colour-word	163.51 (240.35)*	*t*(6) = 2.23	0.059	0.84	Large
	Counting	93.09 (112.65)*				
Cognitive interference of wave V amplitude	Colour-word	-0.02 (0.42)**	*Z*(6) = -1.104	0.32	0.42	Medium
	Counting	0.38 (0.39)**				
Cognitive interference of wave V latency	Colour-word	0.154 (0.244)*	*t*(6) = 1.10	0.32	0.42	Small–medium
	Counting	0.016 (0.184)*				

Note: * = value for Mean (SD); ** = value for Median (IQR); SD = standard deviation; IQR = interquartile range

DISCUSSION

The present study examined differences in behavioural performance on Stroop tasks and electrophysiological responses measured through ABR to assess the effects of cognitive interference elicited by two types of Stroop tasks: the colour-word Stroop task and the counting Stroop task. Overall, the findings revealed no significant differences in Stroop interference (measured by percentage of correct responses and response time) or cognitive interference (measured by ABR wave V amplitude and latency) between the two types of Stroop task in either age group. However, medium to large effect sizes observed in several comparisons, particularly in Stroop interference of response time and cognitive interference of wave V amplitude among children, suggest that task-related variations in cognitive processing demands, especially in inhibitory control, may exist.

Effects on Behavioural Performance of Stroop Task

The absence of statistically significant differences in Stroop interference between the colour-word and counting Stroop tasks suggests that both tasks engage similar inhibitory control mechanisms. Conceptually, the two Stroop paradigms share comparable cognitive demands, as in both tasks, participants are required to attend to a physical dimension of the stimulus (i.e., colour in the colour-word Stroop task and quantity in the counting Stroop task), while ignoring its symbolic dimension (i.e., word meaning in the colour-word Stroop task and digit identity in the counting Stroop task) (Hershman et al., 2024).

At the neurocognitive level, findings from neuroimaging studies have demonstrated that both types of Stroop tasks activate overlapping neural networks, particularly the prefrontal cortex, anterior cingulate cortex (ACC), and parietal cortex (Bush et al., 1998, 2006; Christensen et al., 2011; Freund et al., 2021; Müller et al., 2024). The ACC primarily detects and evaluates cognitive conflict, while the prefrontal cortex implements top-down inhibitory control to resolve such conflict (Ehlis et al., 2024; Shenhav et al., 2013). Meanwhile, the parietal cortex contributes to maintaining sustained attention toward task-relevant features of the stimulus in the presence of more salient but irrelevant features of the stimulus. Activation of parietal cortex during both Stroop tasks has also been associated with the processing of numerical quantity in the counting Stroop task and with lexical or word-reading processes in the colour-word Stroop task (Adleman et al., 2002; Freitas et. al., 2023; Tang et al., 2009). These overlapping neural substrates likely account for the comparable behavioural performance observed between the two types of Stroop task in the present study.

Among children, however, a trend toward longer response times during the incongruent condition of the colour-word Stroop task compared to the counting Stroop task (resulting in greater Stroop interference of response time) was observed. Although this difference did not reach statistical significance, the large effect size suggests a potentially meaningful variation in cognitive load. Upon closer examination, this pattern may reflect a speed-accuracy trade-off (Forte et al., 2024). Since Stroop interference in accuracy (percentage of correct responses) did not differ significantly between the two tasks, it is plausible that children prioritised accuracy over speed when performing the more cognitively demanding colour-word Stroop task. This behaviour suggests that children exercised greater caution and inhibitory control to maintain accuracy, resulting in slower response time. Collectively, these findings imply that the incongruent colour-word Stroop task may impose greater cognitive interference than the incongruent counting Stroop task among developing populations (children).

Effects on Electrophysiological Responses of ABR

Consistent with the behavioural findings, no significant differences were observed in the cognitive interference measurement of ABR between the colour-word and counting Stroop tasks. As both tasks likely impose comparable cognitive demands and engage similar inhibitory control mechanisms, the degree of cognitive interference elicited by each task appears relatively equivalent. Consequently, the level of neural suppression reflected in ABRs was also comparable across both tasks, as evidenced by the absence of statistically significant differences in the cognitive interference of wave V amplitude and latency among both children and adults.

A closer examination of the cognitive interference of wave V amplitude among children, however, revealed a medium effect size, suggesting a potential trend worth noting. Specifically, the colour-word Stroop task demonstrated a more negative pattern of cognitive interference, whereas the counting Stroop task exhibited a positive pattern. The negative interference pattern in the colour-word Stroop task indicates that the wave V amplitude during the incongruent condition was smaller than that observed during the neutral condition. This reduction in amplitude aligns with previous literature suggesting that increased cognitive load and conflict, such as those experienced during the incongruent Stroop condition, can activate the sensory gating mechanism that suppresses neural responsiveness to the concurrent auditory stimuli

presented during the ABR test (Brännström et al., 2018; Dzulkarnain et al., 2021; Shahrudin et al., 2023; Sörqvist et al., 2012). Such suppression, manifested as a decrease in ABR wave V amplitude, likely reflects the reallocation of attentional and processing resources toward resolving cognitive conflict rather than processing auditory input (Brännström et al., 2018; Dzulkarnain et al., 2021; Shahrudin et al., 2023; Sörqvist et al., 2012).

In contrast, the positive interference pattern observed in the counting Stroop task, where wave V amplitude was slightly larger during the incongruent condition, may reflect an absence of substantial neural suppression. This could suggest that the level of cognitive interference induced by the counting Stroop task was not sufficient to strongly engage the sensory gating system. In other words, the neural resources required to resolve conflict in the counting Stroop task may not have been extensive enough to interfere with concurrent auditory processing, resulting in a relatively stable or slightly enhanced ABR response. This interpretation is consistent with prior research suggesting that early auditory processing, as reflected in the ABR, is relatively resilient to mild cognitive load (Brännström et al., 2018).

Taken together, these findings suggest that although both Stroop tasks engage similar inhibitory control mechanisms behaviourally, the colour-word Stroop task may impose a somewhat greater cognitive load at the neural level. This is reflected in the direction of change in wave V amplitude, even though this difference did not reach statistical significance.

Developmental Differences in the Effects of Different Types of Stroop Tasks

Although no statistically significant differences were found in any measurement parameters between the colour-word and counting Stroop tasks among both children and adults, a consistent pattern of medium to large effect sizes was observed only among children. Specifically, the medium and larger effect sizes in Stroop interference of response time and cognitive interference of wave V amplitude, respectively, were evident in the child group. These findings suggest that cognitive interference effects are more pronounced during the colour-word Stroop task among children, as reflected by longer response times and reduced wave V amplitude during the incongruent condition. This pattern implies that children are more susceptible to the higher cognitive load and conflicts induced by the colour-word Stroop task compared to adults. This interpretation aligns with previous studies reporting that cognitive interference effects are much stronger in children, who generally exhibit less efficient response inhibition than adults (Ménétré & Laganaro, 2023).

Several developmental factors may account for these differences. Previous research indicates that the ability to detect cognitive conflict (conflict detection) typically matures around the age of 10 (Ménétré & Laganaro, 2023). However, the ability to resolve such conflicts (conflict resolution) continues to develop through late childhood and adolescence and reaches adult-like levels only in late adolescence or early adulthood (Forte et al., 2024; Jongen & Jonkman, 2008; Ménétré & Laganaro, 2023). Neuroimaging evidence further supports this view, showing that the maturation of conflict resolution abilities parallels the structural and functional brain development, particularly involving the prefrontal cortex (Luna et al., 2010), which is among the last brain regions to mature (Ferguson et al., 2021; Forte et al., 2024; Güntekin et al., 2023). Besides, during the developmental period, neural networks undergo gradual structural changes, including synaptic pruning, grey matter thinning, and myelination (Jongen & Jonkman, 2008), which collectively contribute to improved efficiency in inhibitory control function.

Moreover, although the neural networks involved in inhibitory control in children and adults share common regions such as the ACC and the prefrontal cortex, an activation of additional brain regions was identified in children, mainly in the orbito-frontal regions. The requirement of these additional brain areas may reflect compensatory mechanisms, indicating that children require greater neural resources to achieve successful inhibition (Ménétré & Laganaro, 2023). This is consistent with the notion that the more practiced or automated a cognitive ability becomes, the fewer neural regions are required for its execution (Yotsumoto et al., 2008).

CONCLUSION

The present study investigated the effects of cognitive interference elicited by two types of Stroop tasks, the colour-word and counting Stroop tasks, on behavioural performance and electrophysiological responses measured via ABR among children and adults. Although no statistically significant differences were observed in Stroop interference or in the cognitive interference of ABR between the two tasks, the pattern of results and the medium to large effect sizes observed among children suggest meaningful variations in task-related cognitive processing demands. Specifically, the colour-word Stroop task appeared to impose greater cognitive interference, thereby eliciting stronger inhibitory control, as reflected by longer response times and reduced ABR wave V amplitude during the incongruent condition. These findings suggest

that, although both Stroop paradigms engage overlapping inhibitory control mechanisms, the colour-word Stroop task may induce stronger cognitive conflict, particularly during developmental stages when executive functions are still maturing. From a clinical perspective, the greater susceptibility of children to cognitive interference induced by the colour-word Stroop task, and its corresponding effect on ABR, suggests that this paradigm could serve as a sensitive electrophysiological marker for assessing inhibitory control abilities in children. Such an approach may aid in the early identification of children with atypical inhibitory control, including those with auditory processing disorder (APD) or related neurodevelopmental conditions.

ACKNOWLEDGEMENT

This research was funded by the Ministry of Higher Education, Malaysia through the Fundamental Research Grant Scheme (FRGS) (Project ID: FRGS/1/2024/SKK06/UIAM/02/4).

REFERENCES

Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H., & Reiss, A. L. (2002). A developmental fMRI study of the Stroop color-word task. *NeuroImage, 16*(1), 61–75. https://doi.org/10.1006/nimg.2001.1046

Brännström, K. J., Wilson, W. J., & Waechter, S. (2018). Increasing cognitive interference modulates the amplitude of the auditory brainstem response. *Journal of the American Academy of Audiology, 29*(06), 512–519. https://doi.org/10.3766/jaaa.17003

British Society of Audiology (BSA) (2019). *Recommended procedure auditory brainstem response (ABR) testing in babies.* Bathgate: British Society of Audiology.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: an interference task specialized for functional neuroimaging--validation study with functional MRI. *Human Brain Mapping, 6*(4), 270–282. https://doi.org/10.1002/(SICI)1097-0193(1998)6:4%3C270::AID-HBM6%3E3.0.CO;2-0

Bush, G., Whalen, P. J., Shin, L. M., & Rauch, S. L. (2006). The counting Stroop: a cognitive interference task. *Nature Protocols, 1*(1), 230–233. https://doi.org/10.1038/nprot.2006.35

Christensen, T. A., Lockwood, J. L., Almryde, K. R., & Plante, E. (2011). Neural substrates of attentive listening assessed with a novel auditory Stroop task. *Frontiers in Human Neuroscience, 4*(236), 1-12. https://doi.org/10.3389/fnhum.2010.00236

Dzulkarnain, A. A. A., Tengku Zam Zam, T. Z. H., Azed, Z., Zuri, M. I. R., & Sulaiman, N. H. (2014). Effects of electrode positions on tone-burst-evoked auditory brainstem responses (ABR) in Humans. *Middle-East Journal of Scientific Research, 21*(8), 1180-1187. https://doi.org/10.5829/idosi.mejsr.2014.21.08.84167

Dzulkarnain, A. A. A., Azizi, A. K., & Sulaiman, N. H. (2020). Auditory sensory gating in Huffaz using an auditory brainstem response with a psychological task: A preliminary investigation. *Journal of Taibah University Medical Sciences, 15*(6), 495–501. https://doi.org/10.1016/j.jtumed.2020.08.007

Dzulkarnain, A. A. A., Sulaiman, N. H., Shahrudin, F. A., Sidek, S. N., Md. Yusof, H., & Khalid, M. (2021). Auditory brainstem response with cognitive interference in normal and autism spectrum disorder children - Understanding the auditory sensory gating mechanism. In *IntechOpen*. IntechOpen. https://doi.org/10.5772/intechopen.101135

Ehlis, A.-C., Zarantonello, L., Haeussinfer, F. B., Rohe, T., Rosenbaum, D., Fallgatter, A. J., & Maier, M. J. (2024). The DLPFC is centrally involved in resolving Stroop conflicts, suppressing distracting sensory input within the auditory and visual system. *Frontiers in Psychology, 15*, 1427455. https://doi.org/10.3389/fpsyg.2024.1427455

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports, 11*(1), 1382. https://doi.org/10.1038/s41598-020-80866-1

Forte, G., Troisi, G., Favieri, F., & Casagrande, M. (2024). Inhibition changes across the lifespan: experimental evidence from the Stroop task. *BMC Psychology*, *12*(1), 366. https://doi.org/10.1186/s40359-024-01844-0

Freitas, P. M., Haase, V. G., & Wood, G. M. (2023). The neural correlates of interference effects of numerical Stroop task: An ALE meta-analysis and connectometry. *Progress in brain research, 282*, 71–93. https://doi.org/10.1016/bs.pbr.2023.10.003

Freund, M. C., Bugg, J. M., & Braver, T. S. (2021). A representational similarity analysis of cognitive control

during color-word Stroop. *The Journal of neuroscience: the official journal of the Society for Neuroscience, 41*(35), 7388–7402. https://doi.org/10.1523/JNEUROSCI.2956-20.2021

Güntekin, B., Alptekin, S., Yıldırım, E., Aktürk, T., Uzunlar, H., Çalışoğlu, P., Ada, F. E., Atay, E., & Ceran, Ö. (2023). Immature event-related alpha dynamics in children compared with the young adults during inhibition shown by day-night Stroop task. *Frontiers in human neuroscience, 17*, 1218559. https://doi.org/10.3389/fnhum.2023.1218559

Hershman, R., Keha, E., Sapir, A., Weiss, E. M., Henik, A., & Kaufmann, L. (2024). Evidence for two types of task conflict in a color-digit Stroop task. *Journal of cognition, 7*(1), 54. https://doi.org/10.5334/joc.386

Jones, L. A., Hills, P. J., Dick, K. M., Jones, S. P., & Bright, P. (2016). Cognitive mechanisms associated with auditory sensory gating. *Brain and cognition, 102*, 33–45. https://doi.org/10.1016/j.bandc.2015.12.005

Jongen, E. M. M., & Jonkman, L. M. (2008). The developmental pattern of stimulus and response interference in a colour-object Stroop task: an ERP study. *BMC Neuroscience, 9, 82.* https://doi.org/10.1186/1471-2202-9-82

Knott, V., Millar, A., & Fisher, D. (2009). Sensory gating and source analysis of the auditory P50 in low and high suppressors. *NeuroImage, 44*(3), 992–1000. https://doi.org/10.1016/j.neuroimage.2008.10.002

Korzyukov, O., Pflieger, M. E., Wagner, M., Bowyer, S. M., Rosburg, T., Sundaresan, K., Elger, C. E., & Boutros, N. N. (2007). Generators of the intracranial P50 response in auditory sensory gating. *NeuroImage, 35*(2), 814–826. https://doi.org/10.1016/j.neuroimage.2006.12.011

Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain and cognition, 72*(1), 101–113. https://doi.org/10.1016/j.bandc.2009.08.005

Magnée, M. J. C. M., Oranje, B., Engeland, H. V., Kahn, R. S., & Kemner, C. (2009). Cross-sensory gating in schizophrenia and autism spectrum disorder: EEG evidence for impaired brain connectivity? *Neuropsychologia, 47*(7), 1728–1732. https://doi.org/10.1016/j.neuropsychologia.2009.02.012

Ménétré, E., & Laganaro, M. (2023). The temporal dynamics of the Stroop effect from childhood to young and older adulthood. *PloS one, 18*(3), e0256003. https://doi.org/10.1371/journal.pone.0256003

Müller, V. I., Cieslik, E. C., Ficco, L., Tyralla, S., Sepehry, A. A., Aziz-Safaie, T., Feng, C., Eickhoff, S. B., & Langner, R. (2024). Not All Stroop-type tasks are alike: Assessing the impact of stimulus material, task design, and cognitive demand via meta-analyses across neuroimaging studies. *Neuropsychology review*, 10.1007/s11065-024-09647-1. Advance online publication. https://doi.org/10.1007/s11065-024-09647-1

Potter, D., Summerfelt, A., Gold, J., & Buchanan, R. W. (2006). Review of clinical correlates of P50 sensory gating abnormalities in patients with schizophrenia. *Schizophrenia bulletin, 32*(4), 692–700. https://doi.org/10.1093/schbul/sbj050

Scarpina, F., & Tagini, S. (2017). The Stroop color and word test. *Frontiers in psychology, 8*, 557. https://doi.org/10.3389/fpsyg.2017.00557

Schramm, M., Fjaellingsdal, T. G., Aslan, B., Jung, P., Lux, S., Schulze, M., & Philipsen, A. (2023). Electrophysiological evidence for increased auditory crossmodal activity in adult ADHD. *Frontiers in neuroscience, 17,* 1227767. https://doi.org/10.3389/fnins.2023.1227767

Shahrudin, F. A., Dzulkarnain, A. A. A., Chahed, N., Ahmad Basri, N., Jamal, F. N. Sidek, S. I., Md. Yusof, H., & Abd. Hamid, S. R. (2023). Auditory sensory gating in auditory brainstem response (ABR) with psychological task in autism spectrum disorder (ASD) child: A case study. *International Journal of Allied Health Sciences, 7*(5), 512-527, https://doi.org/10.31436/ijahs.v7i5.878

Sharif, S. (2019). *Statistics for nonstatisticians.* Khazanah Darulaman.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron, 79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Sörqvist, P., Stenfelt, S., & Rönnberg, J. (2012). Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: toward a unified view of attention. *Journal of cognitive neuroscience*, *24*(11), 2147–2154. https://doi.org/10.1162/jocn_a_00275

Tang, J., Critchley, H. D., Glaser, D. E., Dolan, R. J., & Butterworth, B. (2009). Imaging informational conflict: A functional magnetic resonance imaging study of numerical Stroop. *Journal of cognitive neuroscience, 18*(12), 2049-2062. https://doi.org/10.1162/jocn.2006.18.12.2049

Villarreal, G., & Hunter, M. A. (2015). Auditory M50 sensory gating deficit and superior temporal gyrus pathology in posttraumatic stress disorder. In *Springer eBooks* (pp. 1–11). Springer Nature. https://doi.org/10.1007/978-3-319-08613-2_100-1

Yotsumoto, Y., Watanabe, T., & Sasaki, Y. (2008). Different dynamics of performance and brain activation in the time course of perceptual learning. *Neuron, 57*(6), 827–833. https://doi.org/10.1016/j.neuron.2008.02.0

ORIGINAL ARTICLE

Virtual Reality-Based Vestibular Rehabilitation Therapy: A Scoping Review

Megat Ikmal Megat Abdul Karim @ Rosmady [1], Nurul Syarida Mohd Sakeri[1], Saiful Adli Jamaluddin[1]

[1]Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Pahang, Malaysia

Keywords:
Scoping review; virtual-reality; vestibular rehabilitation

ABSTRACT

Introduction: Virtual reality (VR) technologies have been increasingly developed to support vestibular rehabilitation therapy by providing immersive and interactive environments that enhance patient engagement, improve therapeutic outcomes while ensuring patient safety. This scoping review aims to identify and compare the various types of VR-based vestibular rehabilitation approaches currently available. **Methods**: A scoping review was conducted following the Joanna Briggs Institute (JBI) framework to explore and compare distinct VR-based vestibular rehabilitation methods. Studies were included if they utilized VR in vestibular rehabilitation, were published in English between 2012 and 2022, and were available in full text. Searches were performed across electronic databases including PubMed, ScienceDirect, and ProQuest. Relevant data from the selected studies were summarized descriptively. **Results**: A total of nine studies met the inclusion criteria, comprising one observational study and eight experimental studies. Three primary categories of VR devices were identified: head-mounted displays, force plate systems, and hybrid VR setups. These devices were used to deliver vestibular rehabilitation interventions across various protocols and settings. **Conclusion**: All included studies reported that VR-based vestibular rehabilitation was effective and well-tolerated, despite notable differences in intervention procedures and outcome assessments. A systematic review is recommended to further evaluate the comparative effectiveness of VR-based versus conventional vestibular rehabilitation therapies.

INTRODUCTION

Vestibular Rehabilitation Therapy (VRT) is a specific exercises designed to improve symptoms (such as dizziness, vertigo, imbalance, gaze instability, and difficulties with daily activities) caused by vestibular organ dysfunction. VRT integrates physical exercises with multisensory stimulation to improve balance and reduce symptoms in adults (Alahmari et al., 2014). It is considered a safe and effective therapeutic approach, leveraging central neuroplastic mechanisms, which are adaptation, habituation, and substitution, to enhance postural stability and visual-vestibular integration, particularly in environments with conflicting sensory inputs (Lee et al., 2020). This, in turn, contributes to improved self-confidence and quality of life for affected individuals.

However, conventional VRT often requires patients to perform repetitive exercises for 20–40 minutes daily, which can be perceived as monotonous (Herdman, 2007, as cited in Rosiak et al., 2019). Additionally, patients may struggle to receive adequate feedback when training independently at home (Rosiak et al., 2019). These limitations have prompted interest in virtual reality (VR) technologies as a promising alternative for home-based rehabilitation.

VR is a widely adopted method for creating interactive, computer-generated simulations of real-world environments, using various display formats such as spherical screens, flat monitors, or head-mounted displays (HMDs) (Rosiak et al., 2019). Several VR systems have been applied in vestibular rehabilitation, including the Nintendo Wii Fit (which uses a force plate), treadmill-based simulations in virtual environments, and immersive HMD-based setups. Unlike traditional equipment, these VR platforms incorporate gamified elements that enhance enjoyment, focus, and patient adherence (Song, 2019). Moreover, VR provides real-time feedback and offers a more engaging and enriched therapeutic experience compared to conventional methods (Dara et al., 2012).

Although early VR systems were limited by high costs and space requirements, recent advancements have led to the development of affordable and compact devices—such as the Wii balance board and HMDs—that have demonstrated promising results and are increasingly viable for widespread rehabilitation use (Song, 2019).
While numerous studies have compared conventional and VR-based vestibular rehabilitation, few have focused specifically on the distinct types of VR technologies employed. Although the effectiveness of VR-based interventions has been well-documented (e.g., Alahmari et al., 2014; Meldrum et al., 2015), the diversity of VR

devices—such as head-mounted displays, force plate systems, and hybrid VR setups—warrants closer examination. Therefore, this study aims to identify and compare the distinct types of virtual reality-based vestibular rehabilitation currently available.

MATERIALS AND METHODS

Study Design

A scoping review was conducted for this study to identify and compare the distinct types of virtual reality-based vestibular rehabilitation. This study follows the updated scoping review framework from the Joanna Briggs Institute (JBI), introduced by Peters and colleagues in 2020. Their approach provides a clearer and more structured way to carry out scoping reviews, making the process more comprehensive and reliable. This framework includes; 1. Define objective and research questions, 2. Develop inclusion and exclusion criteria, 3. Data searching, 4. Data selection, 5. Data extraction, and 6. Data analysis.

Criteria of Study Characteristics

Any studies available in full text, written in English, published between 2012 and 2022, and involving VR-based vestibular rehabilitation were included. All study designs that provided a quantitative evaluation of virtual reality-based vestibular rehabilitation programs (e.g., randomized controlled trials, non-randomized controls, cohort studies, case-control studies, cross-sectional studies, and pre-post studies) were considered. The studies also needed to include therapeutic evidence and be specifically designed for patients with central or peripheral vestibular disorders. Studies focusing solely on augmented reality technology, those in the developmental phase of virtual reality-based vestibular rehabilitation, and review articles were excluded.

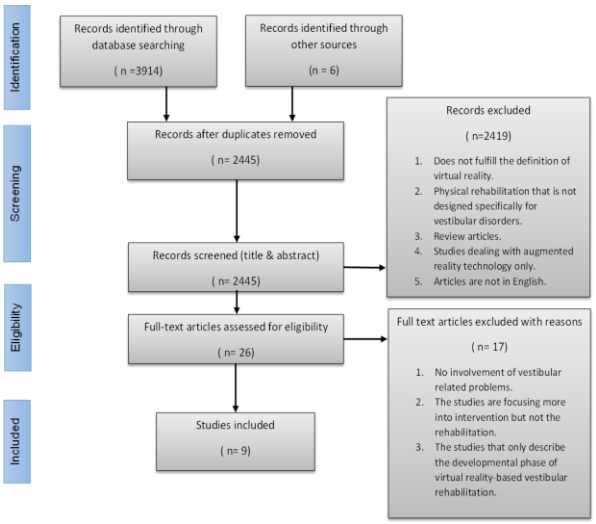


Figure 1: PRISMA flow diagram

Data Searching

Three electronic databases (i.e.; Pubmed, ScienceDirect, and ProQuest) were utilized for the searching of the related studies. The searches were performed using free-text keywords and combining keywords with Boolean operators which are "AND" and "OR" to identify the relevant studies. The search keywords are categorized into three blocks; "virtual reality" (block 1), AND "vestibular" (block 2), AND "rehabilitation therapy" (block 3) (refer to Table 1). A manual search through the reference list of identified articles was also conducted.

Table 1: Search Keyword and Boolean operator. The keyword group is combined by "AND"

Block	Block 1	Block 2	Block 3
	Virtual reality is combined by "OR"	Vestibular is combined by "OR"	Rehabilitation therapy is combined by "OR"
Keywords	Virtual reality Simulation	Vestibular Balance	Rehabilitation therapy Physical therapy Excercise program

Data Selection

The identified articles underwent three stages of screening. First, the main author screened the title and removed any duplicates of studies. The full text of selected articles was obtained and reviewed for eligibility. Next, articles were reviewed independently by two independent reviewers who are the team members of this study. Only eligible articles that met the inclusion criteria were selected. Any disagreement between the two reviewers was discussed and solved by the second co-author. The process of selecting articles was summarized on a PRISMA flow diagram (Figure 1).

RESULTS

A total of 3,914 articles were identified through electronic database searches, with an additional six articles retrieved from reference lists of related studies. After removing duplicates, 2,445 articles remained for title and abstract screening. Of these, 2,419 were excluded, leaving 26 articles for full-text review. Ultimately, nine articles met the inclusion criteria and were included in the final analysis. Figure 1 illustrates the search and selection process, along with reasons for exclusion.

Study Characteristics

Among the nine studies reviewed, one was an observational prospective study, while the remaining eight were experimental in design (randomized controlled trials and pre-post studies). Specifically, seven studies employed randomized controlled trials (Garcia et al., 2013; Meldrum et al., 2015; Eftekharsadat et al., 2015; Anson et al., 2018; Viziano et al., 2019; Stankiewicz et al., 2020; Sessoms et al., 2021). The other two studies included one pre-post study (Alahmari et al., 2014) and one prospective observational study (Rosiak et al., 2019).

In total, 378 subjects were involved in the selected studies. Number of subjects ranged from 20 to 71 individuals with number of subjects for each group range between 10 and 36 individuals. The age range of the subjects were from 18 to 92 years. Due to a lack of information in the texts, other demographic data could not be extracted from the studies examined. The following data is based solely on patients with vestibular disorders who have undergone virtual reality rehabilitation. There are five studies (56%) that focused on the subjects with peripheral vestibular disorders only (Garcia et al., 2013; Meldrum et al., 2015; Viziano et al., 2019; Rosiak et al., 2019; Stankiewicz et al.,2020), two studies (22%) that involved the subjects with central vestibular disorders (Eftekharsadat et al., 2015; Sessoms et al., 2021), and the other two (22%) of the reviewed studies included both subjects with peripheral and central vestibular disorders (Alahmari et al., 2014; Anson et al., 2018).

Types of Virtual Reality Devices

Across these studies, three main types of devices were used to deliver virtual reality: head-mounted devices (HMD) (Viziano et al., 2019; Stankiewicz et al., 2020), force plate systems (Alahmari et al., 2014; Meldrum et al., 2015; Eftekharsadat et al., 2015; Anson et al., 2018; Sessoms et al., 2021), and hybrid VR devices (Garcia et al., 2013; Rosiak et al., 2019). Of the nine studies, five studies (56%) utilized force plate devices, while head-mounted and hybrid VR devices each accounted for two studies (22% each). Notably, among the studies using force plate technology, three different systems were employed: a treadmill (Alahmari et al., 2014; Anson et al., 2018; Sessoms et al., 2021), the Biodex Balance System SD (Eftekharsadat et al., 2015), and the Nintendo Wii Fit Plus (Meldrum et al., 2015).

Virtual Reality Technologies and Protocols in Rehabilitation Approaches

Table 2 presents a comparative overview of various studies employing virtual reality (VR) devices for rehabilitation. It categorizes the devices into head-mounted, force plate, and hybrid systems, each offering unique features and therapeutic approaches. Head-mounted devices, such as VR goggles and cockpit-style games, focus on immersive visual experiences and are often used in both clinical and home settings. Force plate devices, including systems like the Nintendo Wii Fit Plus and Biodex Balance System, emphasize balance training through interactive feedback and motion tracking. Hybrid devices combine multiple technologies—such as force platforms, motion sensors, and VR goggles—to deliver comprehensive rehabilitation programs that target postural control and sensory integration.

Across the studies, rehabilitation durations vary from short daily sessions to extended weekly programs over several weeks. Procedures range from supervised clinical exercises to home-based protocols, often involving gamified tasks to enhance engagement. The features of each system are tailored to simulate real-world scenarios (e.g., grocery shopping or rollercoaster rides) or provide visual feedback for motor control. Overall, the table highlights the diversity in VR rehabilitation strategies, showcasing how different technologies and protocols are adapted to meet specific therapeutic goals.

Table 2. Summary of virtual reality-based rehabilitation studies, outlining types of VR equipment, therapy duration, features, and procedures.

Author(s) Year	Types of virtual reality device	Duration of rehabilitation	Features	Procedure
Viziano et al. (2019)	Head-mounted device	(i) Supervised session in the clinic: twice a week for 4 weeks for 30-45 min. (ii) At home: twice daily for 30-40 min.	The Track Speed Racing 3D game featured a point-of-view race in which the car was handled from the cockpit by tilting the head to the left and right to avoid drifting off the road and to accomplish all of the missions before the lap was completed.	Each patient was instructed to sit on a chair or sofa and complete the gaming protocol uninterruptedly.
Stankiewicz et al. (2020)	Head-mounted device: New generation VR goggles based on Google Cardboard platform.	Five therapy sessions for five consecutives days with two sessions of 5-minutes with 5-minute intervals.	The VR goggles were utilized to simulate a rollercoaster ride.	Patients participated in the exercises under medical supervision. They will experience the simulation of a rollercoaster ride through the VR goggles.
Alahmari et al. (2014)	Force plate device: Treadmill and virtual grocery store displayed in an immersive environment	1 session per week for 6 weeks, each session lasted for 1 hour.	Grocery store scene was displayed on three screens that surrounded the subject in a full-field immersive visual environment. The store contained 16 aisles and had eight levels of visual complexity.	Patients are required to walk on a treadmill as if they were in a virtual grocery store, and traveling through the aisles during each treatment session. While ambulating, the therapists will instruct the patients to locate products on the shelf, and the patients will verbally answer when they located the product.
Meldrum et al. (2015)	Force plate device: Nintendo Wii Fit Plus (NWFP)	15-minute daily for 6 weeks	The NWFP combines an accelerometer and a force plate (Wii Balance Board) to provide visual and auditory feedback of subjects' centre of pressure during virtual reality exercises and games.	The subjects were loaned one NWFP for use at home along with a standardized weekly booklet that will detail their exercises and include an adherence diary. The subjects performed a daily rehabilitation programme.
Eftekharsad at et al. (2015)	Force plate device: The Biodex Balance System SD	Twice a week for 12 weeks, each session lasted for 20 minutes	The system simulates specific movement patterns or strategies by placing marks on specified positions on the screen grid.	The subjects attempted to touch targets nine times in each session using an on-screen cursor that was manipulated by the subjects' legs on the device's platform. The stability of the platform was set to 6 (moderately stable).
Anson et al. (2018)	Force plate device: Trunk motion visual feedback	3 times per week for four weeks, each session lasted for 30 minutes	The VFB device featured 10 rings of the bull's-eye that was one inch wide and corresponded to one inch of physical space on the treadmill. Two	The subjects were introduced on how to interact with both forms of trunk motion VFB during the first training

	(VFB) during treadmill walking		webcams tracked the 3-D position of the three markers (at the navel, and each shoulder)	session. They were provided a brief training session and displayed the ability to keep the cursor "as close to the center of the bull's-eye as possible," with minimal displacement or angular deviations.
Sessoms et al. (2021)	Force plate device: CAREN system that comprises of an instrumented treadmill	30 minutes twice a week over a 6-week period	The system comprises of an instrumented treadmill embedded on a platform that can move in six degrees of freedom. The platform is surrounded by a panoramic screen that immerses participants in a VR scenario.	The subject involved in a static stance by shifting the centre of mass to the limits of stability and facilitating weight shifting. Then, they engaged with the pitch and roll of the motion platform changes during dynamic tasks like walking to follow the virtual terrain.
Garcia et al. (2013)	Hybrid VR devices: A Balance Rehabilitation Unit (BRU)	45 minutes twice a week for a total of 12 sessions	The equipment includes a computer with the test's program, a safety metal frame, protection support with straps and belts, a force platform, virtual reality goggles, an accelerometer, and foam cushions. Three modules were included: posturography, body balance rehabilitation, and postural training games (PTG).	Patients were instructed to stand in an orthostatic position without moving their upper limbs, heels, or feet for 60 seconds in order to assess each of the ten sensory conditions. A foam cushion of medium density was used in the third test condition. Virtual reality goggles were used from the fourth to the tenth condition.
Rosiak et al. (2019)	Hybrid VR devices	Ten 30-minute training sessions over the course of two weeks	A hybrid VR equipment comprising a force plate, an upper-body motion sensor, and a central unit with a flat-screen display. The software comprises 8 balance exercises with 28 difficulty levels that gradually escalate from a predetermined starting point based on the patient's performance.	Patients conducted a set of exercises coordinating upper body movement and maintaining the centre of pressure (COP) in a pre-set range or shifting the COP towards indicated positions while standing 1.5–2 metres from the display on the force plate. The exercise protocol comprised of two tasks: eliminating Meteorites from the outer space and placing hats on hangers with matching colours.

DISCUSSION

This scoping review indicates that the implementation of virtual reality (VR) in vestibular rehabilitation is a promising approach, supported by an increasing number of studies exploring various VR devices. Across the reviewed studies, patients consistently demonstrated symptom improvement, with vertigo-related handicap scores generally shifting from moderate to mild following VR-based interventions. For example, Viziano et al. (2019) reported significant improvements in vestibulo-ocular reflex, static posture, and functional performance during follow-up assessments in patients with unilateral hypofunction using head-mounted display (HMD) technology. Similarly, other studies documented enhanced balance control and gait stability when VR

systems incorporated force plate technology or hybrid VR devices, offering immersive environments that may increase patient engagement while maintaining safety compared to conventional settings. These findings suggest that VR-based rehabilitation not only provides an interactive experience but also delivers measurable functional benefits. Nevertheless, further research is needed to standardize protocols and determine the most effective and cost-efficient strategies for clinical implementation.

Sessoms et al. (2021) highlighted the advantages of incorporating a treadmill into virtual reality (VR)-based vestibular rehabilitation, as it enables objective analysis of functional responses such as gait mechanics and postural control. In their study, participants demonstrated measurable improvements in gait symmetry and stability, which provided clinicians with actionable data to tailor training to individual patient needs. Additionally, VR treadmill training allowed subjects to practice life-like scenarios in a safe environment, which was associated with increased confidence in returning to daily activities. VR devices also offer feedback systems—such as force plate technology, head-mounted displays (HMDs), and hybrid VR platforms—that facilitate motor learning. Schmick and Lee (2005) reported that feedback combined with repetitive practice can significantly enhance motor learning, supporting the theoretical basis for VR interventions. These features present substantial opportunities for therapists to monitor progress and optimize patient outcomes. Despite these promising findings, further research is needed to define optimal training parameters, establish standardized protocols, and identify cost-effective strategies for implementing VR-based vestibular rehabilitation.

The outcome measures used across the analyzed studies varied considerably, encompassing both self-report and performance-based assessments. This variability highlights the need for more rigorous and standardized evaluation methods in future research. For instance, the Dizziness Handicap Inventory (DHI), a widely recognized and easy-to-administer tool for assessing vestibular disorder-related handicaps (Jacobson & Newman, 1990), was not applied consistently; only five studies utilized DHI (Garcia et al., 2013; Alahmari et al., 2014; Eftekharsadat et al., 2015; Viziano et al., 2019; Sessoms et al., 2021). While multiple scales and questionnaires exist, adopting standardized tools such as the DHI would enhance comparability across studies. Additionally, the Vestibular Rehabilitation Benefits Questionnaire (VRBQ) may be considered, as it is specifically designed to measure changes in symptoms and quality of life between rehabilitation sessions (Morris et al., 2008). Establishing consensus on core outcome measures will be essential to improve the rigor and interpretability of future VR-based vestibular rehabilitation research.

From a practical perspective, the amount of time spent in virtual reality-based training and the number of sessions is the most crucial component in increasing the therapeutic effect. It is noteworthy that all the selected studies implemented variety of rehabilitation duration per session which the minimum recorded was 15 minutes, up until 60 minutes per session. Surprisingly, Bergeron et al. (2015) reported that, current meta-analysis appeared to show that the amount of time spent in VR-based therapy was more important than the number of sessions in determining average efficiency. According to the findings, a minimum of 120–150 minutes of exposure time is required to identify a meaningful improvement for the patient. These findings suggest that longer sessions, even with fewer appointments, may be both effective and convenient. However, session duration should be tailored to the patient's physical condition and tolerance to ensure safety and maximize benefit.

One of the strengths of this review is that the studies addressed at a wide range of patients with vestibular disorders which includes peripheral, central, or mixed that are similar to those seen in clinical settings. Virtual reality environments are incredibly beneficial for a variety of diseases. However, one of the possible limitations of implementing VR-based vestibular rehabilitation is the related cost of utilizing such protocols as clinical tools. This could be a major issue against the execution of virtual reality settings in clinical settings if there are crucial attempts of cost reductions in the health care system (Bergeron et al., 2015). Nevertheless, very positive outcomes have been demonstrated with affordable prices such as a Nintendo Wii which was studied by Meldrum et al. (2015). Another limitation that may limit the execution of virtual reality-based vestibular rehabilitation is cybersickness. This condition may occur if there is an exposure to interactive virtual environments that can cause discomfort during or after the session due to unnatural and conflicting multisensory stimuli (Bonalo et al., 2009; Nolin et al., 2012; Kiryu & So, 2007 as cited in Bergeron et al., 2015). According to Nolin et al. (2012), motion sickness-like, including nausea, vomiting, headache, somnolence, loss of balance, and altered eye-hand coordination are among the symptoms reported by the subjects.

Limitation of the study

There are several limitations noted in this review. In terms

of inclusion criteria, it was restricted only to include studies that were published in English and within the last 10 years, which was from 2012 to 2022 as it was aimed to review the most recent technology of virtual reality devices that have been utilized in vestibular rehabilitation therapy. Other languages especially from non-English speaking countries may be considered in future review studies whereby there are possibilities of other types of virtual reality-based vestibular rehabilitation therapy.

CONCLUSION

There are nine studies discovered with distinct types of virtual reality devices and all the rehabilitation procedures demonstrated the promising potential of virtual reality-based rehabilitation for vestibular disorders. Despite major variances in procedure and result evaluation, all studies showed that virtual reality-based rehabilitation methods were effective and well-tolerated. The overall time spent in virtual reality training is the most important parameter for predicting treatment success and the amount of symptom reduction. To sum up, some recommendations for future research are proposed, including standardizing intervention protocols and evaluation tools, documenting side effects, and determining if virtual reality-based rehabilitation should be integrated with conventional rehabilitation. A systematic review is recommended to examine the effectiveness of virtual reality-based vestibular rehabilitation therapy compared to conventional vestibular rehabilitation.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

Alahmari, K. A., Sparto, P. J., Marchetti, G. F., Redfern, M. S., Furman, J. M., & Whitney, S. L. (2014). Comparison of virtual reality based therapy with customized vestibular physical therapy for the treatment of vestibular disorders. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(2), 389–399. https://doi.org/10.1109/TNSRE.2013.2294904

Anson, E., Ma, L., Meetam, T., Thompson, E., Rathore, R., Dean, V., & Jeka, J. (2018). Trunk motion visual feedback during walking improves dynamic balance in older adults: Assessor blinded randomized controlled trial. *Gait and Posture*, *62*(March), 342–348. https://doi.org/10.1016/j.gaitpost.2018.03.044

Bergeron, M., Lortie, C. L., & Guitton, M. J. (2015). Use of Virtual Reality Tools for Vestibular Disorders Rehabilitation: A Comprehensive Analysis. *Advances in Medicine*, *2015*, 1–9. https://doi.org/10.1155/2015/916735

Dara, M., Susan, H., Roisin, M., Deirdre, M., Douglas, D., Kareena, M., Helen, F., Stephen, H., Ronan, C., & Rory, M. W. (2012). Effectiveness of conventional versus virtual reality based vestibular rehabilitation in the treatment of dizziness, gait and balance impairment in adults with unilateral peripheral vestibular loss: A randomised controlled trial. *BMC Ear, Nose and Throat Disorders*, *12*(1), 1–8. https://doi.org/10.1186/1472-6815-12-3

Eftekharsadat, B., Babaei-Ghazani, A., Mohammadzadeh, M., Talebi, M., Eslamian, F., & Azari, E. (2015). Effect of virtual reality-based balance training in multiple sclerosis. *Neurological Research*, *37*(6), 539–544. https://doi.org/10.1179/1743132815Y.0000000013

Garcia, A. P., Ganança, M. M., Cusin, F. S., Tomaz, A., Ganança, F. F., & Caovilla, H. H. (2013). Reabilitação vestibular com realidade virtual na doença de Ménière. *Brazilian Journal of Otorhinolaryngology*, *79*(3), 366–374. https://doi.org/10.5935/1808-8694.20130064

Jacobson, G. P., & Newman, C.W., "The development of the Dizziness Handicap Inventory," *Archives of Otolaryngology—Head and Neck Surgery*, vol. 116, no. 4, pp. 424–427, 1990.

Lee, S., Hong, M., Kim, S., & Choi, S. J. (2020). Effect analysis of virtual-reality vestibular rehabilitation based on eye-tracking. *KSII Transactions on Internet and Information Systems*, *14*(2), 826–840. https://doi.org/10.3837/tiis.2020.02.020

Meldrum, D., Herdman, S., Vance, R., Murray, D., Malone, K., Duffy, D., Glennon, A., & McConn-Walsh, R. (2015). Effectiveness of Conventional Versus Virtual Reality-Based Balance Exercises in Vestibular Rehabilitation for Unilateral Peripheral Vestibular Loss: Results of a Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation*, *96*(7), 1319-1328.e1. https://doi.org/10.1016/j.apmr.2015.02.032

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G (2009). *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ, 339(jul21 1), b2535–b2535.* doi:10.1136/bmj.b2535

Morris, Anna E., Lutman, Mark E., Yardley, Lucy (2008). *Measuring outcome from Vestibular Rehabilitation, Part I: Qualitative development of a new self-report measure. International Journal of Audiology, 47(4), 169–177.* doi:10.1080/14992020701843129

Peters, M. D. J., Godfrey, C., McInerney, P., Soares, C. B., Parker, D., Khalil, H. 2017 Guidance for the Conduct of JBI Scoping Reviews Chapter 11: Scoping Reviews. *Joanna Briggs Institute Reviewer's Manual*. The Joanna Briggs Institute, 2017. https://reviewersmanual.joannabriggs.org/

Rosiak, O., Krajewski, K., Woszczak, M., & Jozefowicz-Korczynska, M. (2019). Evaluation of the effectiveness of a Virtual Reality-based exercise program for Unilateral Peripheral Vestibular Deficit. *Journal of Vestibular Research: Equilibrium and Orientation*, *28*(5–6), 409–415. https://doi.org/10.3233/VES-180647

Sessoms, P. H., Fraser, J. J., Bodell, D. M., Haluch, K. S., Markham, A. E., Brookfield, J. S., Jameson, J., & Gottshall, K. R. (2021). Clinical effectiveness of virtual reality versus conventional clinic-based vestibular physical therapy on balance and function in active duty service members. A pilot randomized controlled trial. *Virtual Reality*, *Dvbic 2019*. https://doi.org/10.1007/s10055-021-00546-x

Song, J. J. (2019). Virtual reality for vestibular rehabilitation. *Clinical and Experimental Otorhinolaryngology*, *12*(4), 329–330. https://doi.org/10.21053/ceo.2019.00983

Stankiewicz, T., Gujski, M., Niedzielski, A., & Chmielik, L. P. (2020). Virtual reality vestibular rehabilitation in 20 patients with vertigo due to peripheral vestibular dysfunction. *Medical Science Monitor*, *26*, 1–9. https://doi.org/10.12659/MSM.930182

Viziano, A., Micarelli, A., Augimeri, I., Micarelli, D., & Alessandrini, M. (2019). Long-term effects of vestibular rehabilitation and head-mounted gaming task procedure in unilateral vestibular hypofunction: a 12-month follow-up of a randomized controlled trial. *Clinical Rehabilitation*, *33*(1), 24–33. https://doi.org/10.1177/0269215518788598

ORIGINAL ARTICLE

Analysis of Auditory Brainstem Response Test Time Using Narrow Band Level Specific CE-Chirp and Tone Burst in Infants

Amnah Mohamad Noor[1], Ahmad Aidil Arafat Dzulkarnain,[1,2*], Saiful Adli Jamaluddin[1], and Sarah Rahmat[1,2], Siti Rashidah Mohd Zain[3]

[1]Department of Audiology and Speech Language-Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Pahang, Malaysia
[2]Children Health and Wellbeing research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Pahang, Malaysia
[3]Department of Otorhinolaryngology Sultan Ahmad Shah Medical Centre, International Islamic University Malaysia, 25200, Pahang, Malaysia

*Corresponding author: Ahmad Aidil Arafat Dzulkarnain, ahmadaidil@iium.edu.my

Keywords:
Narrow Band Level Specific CE-Chirp; Fmp; auditory brainstem response; hearing threshold estimation; infants.

ABSTRACT

Background and Aims: Despite advancements in auditory brainstem response (ABR) stimuli, the clinical application of narrow-band level-specific CE-Chirp (NB LS CE-Chirp) and the objective F-test multiple points (Fmp) algorithm remains underexplored, particularly in infant populations. These methodologies hold promise for improving the efficiency of ABR assessments and universal newborn hearing screening. This study evaluates the use of NB LS CE-Chirp as a test stimulus and the Fmp statistical algorithm as a stopping criterion to address prolonged ABR test durations. **Subjects and Methods:** Fifty infants (27 males, 23 females) under six months of age were enrolled. ABR recordings were obtained using both NB LS CE-Chirp and tone burst (TB) stimuli at 500, 1000, and 4000 Hz, presented at 70 and 40 dBnHL. Test durations were calculated using two stopping criteria: Fmp ≥ 3.1 and visual detection. **Results:** NB LS CE-Chirp stimuli elicited significantly shorter test durations across all frequency and intensity combinations compared to TB stimuli, under both Fmp ≥ 3.1 and visual detection criteria. Notably, the Fmp ≥ 3.1 algorithm did not consistently yield shorter test times than visual detection for either stimulus type. **Conclusion:** The use of NB LS CE-Chirp stimuli, combined with either visual or Fmp-based detection algorithms, presents a promising strategy for infant hearing assessment. This approach significantly reduces ABR test duration and may enhance the efficiency of newborn hearing screening protocols.

INTRODUCTION

Universal Newborn Hearing Screening (UNHS) programs, implemented both in Malaysia and globally, aim to identify infants with congenital hearing loss at the earliest possible stage. These initiatives are aligned with the recommendations of the Joint Committee on Infant Hearing (JCIH), which advocate for Early Hearing Detection and Intervention (EHDI) protocols. According to EHDI guidelines, newborns should undergo hearing screening within the first day of life or, at the latest, by one month of age. A confirmed diagnosis of hearing loss is expected by two months, followed by initiation of aural rehabilitation before three months of age. This structured timeline facilitates timely audiological follow-up and intervention (Malaysia Ministry of Health Screening Committee, 2014).

In Malaysia, audiological services face several challenges in achieving these benchmarks, particularly in rural and remote regions. Barriers include limited access to newborn hearing screening (NHS) programs, delays in follow-up diagnostics such as auditory brainstem response (ABR) testing, and insufficient coordination among healthcare providers. These delays are often attributed to long waiting lists and the inherently time-intensive nature of ABR procedures (Cain et al., 2020).

Conventional ABR stimuli, such as clicks and tone bursts, typically require extended testing durations and may necessitate multiple sessions. This contributes to reduced appointment availability and prolonged waiting periods in clinical settings. Moreover, successful ABR recordings depend on the infant being in a quiet or sleep state to complete a full sweep cycle, typically around 2000 sweeps, which is critical for effective signal averaging and enhancement of the signal-to-noise ratio (SNR).

* *Corresponding author.*
E-mail address: ahmadaidil@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

Consequently, some patients may require two to three separate visits to complete the assessment (Cain et al., 2020).

NB LS CE-Chirp is a recent advancement in frequency-specific auditory brainstem response (FS ABR) stimulus. This stimulus presentation duration varies with stimulus intensity and is precisely calibrated to account for cochlear traveling wave delays at central frequencies of 500, 1000, 2000, and 4000 Hz, thereby enhancing ABR amplitude elicited via tone burst stimulus; however, the study was conducted in adult population (Megha et al., 2019; Cebulla & Elberling, 2010). In addition, it was observed that amplitude elicited from ABR to tone burst is usually 70% smaller than the amplitude of ABR to click as the tone burst is a FS ABR stimulus encompassing a specific frequency range which elicited less robust overall response (Ferm et al., 2013). Due to its ability to elicit robust neural responses, NB LS CE-Chirp may achieve an adequate SNR more rapidly, potentially reducing overall test duration. In contrast, tone burst stimulus at 500 to 4000 Hz with low to moderate stimulus intensity levels produce less robust and earlier latency of waves (I, II, III and IV), and these early peaks are misinterpreted as wave V. From previous literature, poor quality ABR was produced by tone burst, specifically at 500 Hz, in which the waves are complex with unstable responses for visual detection (Pinto & Matas, 2007).

To our knowledge, no published studies have directly compared ABR acquisition times between NB LS CE-Chirp and conventional tone burst stimuli in infants. Dzulkarnain et al. (2018) studied the effects of different electrode configurations on the NB LS CE-Chirp and tone-burst auditory brainstem response at multiple intensity levels and frequencies in normal hearing adult. The current study aims to compare ABR acquisition times using NB LS CE-Chirp and tone burst stimuli at frequency-specific levels of 500, 1000, and 4000 Hz, employing two stopping criteria using visual detection and Fmp ≥ 3.1 at 70 and 40 dBnHL in infants with normal hearing.

MATERIALS AND METHODS

Participants recruitment

This study received ethical approval from the Kulliyyah Postgraduate and Research Committee (KPGRC), IIUM Research Ethics Committee [IREC 2023-065], Medical Research and Ethics (MREC) National Medical Research Registration [(NMRR) ID-23-01821-J12 (IIR)], and Sultan Ahmad Shah Medical Centre (SASMEC) approval from the Department of Ear, Nose, and Throat (ENT) and Department of Obstetrics and Gynaecology (O & G). Fifty infants (27 males and 23 females) aged below 6 months old with a mean age of 2.55 months (SD = 0.83) met the inclusion criteria of 1) normal ear canal and middle ear conditions, 2) passed the neonatal hearing screening program using distortion product otoacoustic emissions (DPOAE), 3) Type A high-frequency tympanogram suggesting normal middle ear function based on Jerger classification, 4) no history of otological diseases, 5) no upper respiratory tract infection (URTI) symptoms at least two weeks before appointment day, and 6) infants without high-risk factors of hearing loss, according to JCIH (2019). Participants were identified and selected from the Department of Obstetrics and Gynaecology (O&G) and Otorhinolaryngology (ORL) at SASMEC. The initial contact was made by the researchers via telephone or during an appointment, where the study was introduced and an appointment was scheduled. Informed consent was subsequently obtained in person during the scheduled appointment, which was also conducted by the researcher.

Methods

The study was conducted in a sound-treated room at IIUM Hearing and Speech Clinic, Jalan Hospital Campus, Kuantan, Pahang. In this study, ABR measurements were conducted using the Interacoustics Eclipse EP25 system. ABR assessment was performed in a sound-treated room, and the stimulus was given via ER-3A (Etymotic Research) insert earphones. The ABR using NB LS CE-Chirp and tone burst stimuli were tested at three specific frequencies of 500, 1000, and 4000 Hz at 70 and 40 dBnHL using a 33.33 Hz stimulus rate with filter settings 33Hz for high-pass and 3000 Hz for low-pass, and ≤ 40 microvolt artifact rejection.

Subjects need to be in a relaxed state and, if possible, sleep on the bed or on the caretaker's lap during the recording session to reduce muscle artifact at low frequencies. Subjects were under natural sleep without sedation. ABR was recorded using a two-channel function through four electrode locations. Before placing the electrode, the subject was prepared by scrubbing the respective places for electrode placement using NuPrep Skin Preparation electrode gel. Ipsilateral and contralateral electrode montage was used by placing Ambu Neuroline 720 disposable electrodes on high forehead marks as a non-inverting (Cz) electrode, lower forehead marks as a ground electrode (Fpz), inverting electrodes on both mastoids at 70 and 40 dBnHL, see Table 1 (Dzulkarnain et al., 2018). The electrode impedance was kept below 5 kΩ during all measurements, measured using the ABR Interacoustics Eclipse EP25 impedance check function, and the inter-electrode impedance was balanced. These two-channel configurations were used to obtain information from two

distinct locations on the inverting electrode, which are optimal for capturing all ABR peaks (Dzulkarnain et al., 2021).

In this study, the ABR acquisition time was determined using combinations of two stimulus types (NB LS CE-Chirp and tone burst) and two methods of ABR stopping criteria (Fmp and visual detection). The maximum sweep of the test was set to 2500. To quantify the ABR acquisition time, the researcher uses the following steps: i) In this experiment, the researcher "pauses" the signal averaging immediately as the Fmp hits 3.1 (Fmp ≥ 3.1). The Fmp is

Table 1. ABR electrode montages

Electrode Type	Standard Placement Location	Ipsilateral Recording (Example: Left Ear Stimulated)	Contralateral Recording (Example: Left Ear Stimulated)
Non-Inverting (+) / Active	Vertex (Cz) or high forehead (Fpz)	Vertex (Cz) or high forehead (Fpz)	Vertex (Cz) or high forehead (Fpz)
Inverting (-) / Reference	Mastoid or earlobe	Left Mastoid (M1) or Left Earlobe (A1) (same side as stimulus)	Right Mastoid (M2) or Right Earlobe (A2) (opposite side of stimulus)
Ground (⏚)	Low forehead, cheek, or nape of neck	Low forehead, cheek, or nape of neck	Low forehead, cheek, or nape of neck

the statistical F-value for multiple points and provides a better confidence level and detection rate of ABR in less time. Two audiologists confirmed the presence or absence of ABR wave V, recorded and stopped by Fmp ≥ 3.1, using waveform morphology. ii) In terms of visual detection technique, the researcher "paused" the averaging as soon as the ABR waveform occurred and was visually identified. Wave V was plotted, and the number of sweeps at that point was recorded. iii) The ABR acquisition time with various combinations of stimulus types, frequencies, and stopping criteria was documented and analyzed.

Statistical analysis

Statistical analysis was measured using IBM SPSS Statistics for Windows, Version 27.0. (IBM Corp., Armonk, NY, USA). The mean differences of ABR test time among the combinations of two stimuli NB LS CE-Chirps and tone burst), two detection types (Fmp and visual detection) at two intensity levels (70 and 40 dBnHL) and three frequencies (500, 1000, and 4000 Hz) were compared using the Wilcoxon signed-rank test at 95% confidence levels.

RESULTS

Fifty ABRs of normal hearing infants using NB LS CE-Chirp and tone burst were collected in this study. Table 1 shows the median and inter-quartile range of the ABR test time for all combinations using two stimuli (NB LS CE-Chirps and tone burst), two detection types (Fmp and visual detection), two intensity levels (70 and 40 dBnHL), and three frequencies (500, 1000, and 4000 Hz).

Table 3 summarizes the number of ABRs identified by the Fmp detection method at two stimulus intensities (70 and 40 dBnHL) across three test frequencies (500, 1000, and 4000 Hz). Consistent detection between the Fmp technique and audiologist confirmation was observed only at 70 dBnHL for both NB LS CE-Chirp and tone burst stimuli at 1000 Hz. For the remaining frequency and intensity combinations, discrepancies were noted, with Fmp detection showing inconsistencies ranging from 2% to 10%.

Tables 4 and 5 show the p-values and effect sizes obtained from the Wilcoxon signed-rank test, which compares ABR test times across all test combinations.

Comparison of ABR test time between ABR from NB LS CE-Chirp and tone burst

The results from Tables 2, 4 and 5 indicate that ABRs elicited by the NB LS CE-Chirp stimulus had significantly shorter acquisition times compared to those elicited by tone burst stimuli, across both detection techniques, stimulus intensities, and all test frequencies ($p < 0.05$). The only exception was at 4000 Hz for both intensities using the Fmp detection method, where acquisition times for both stimuli were statistically similar ($p \geq 0.05$). Overall, ABRs acquired using the NB Chirp stimulus were faster by approximately 4.51 to 18.02 seconds with Fmp detection, and 9.02 to 15.02 seconds with visual detection, compared to those acquired using tone burst stimuli.

Comparison of ABR test time between Fmp and visual detection

The results from Tables 2, 4, and 5 indicate that ABRs identified using the Fmp detection method were significantly faster than those identified visually by an audiologist in only 5 out of 12 combinations of stimulus intensity, frequency, and type ($p < 0.05$). At 40 dBnHL, visual detection identified ABRs faster than Fmp when elicited by NB Chirp stimuli at 500 Hz and 4000 Hz ($p<0.05$). For the remaining combinations, both detection techniques yielded ABRs with comparable acquisition times.

DISCUSSION

NB LS CE-Chirp versus tone burst

The comparison of ABR acquisition times using NB LS CE-Chirp and tone burst stimuli at 500, 1000, and 4000 Hz, presented at 70 and 40 dBnHL were observed from this study.

Overall, NB LS CE-Chirp elicited shorter ABR acquisition times than tone burst stimulus. This improvement is likely attributable to the enhanced neural synchrony achieved by NB LS CE-Chirp, a subset of broadband LS CE-Chirp. The stimulus design involves precise timing of frequency components, with lower frequencies presented earlier to compensate for cochlear travel delays. This synchronization allows simultaneous activation of cochlear regions, resulting in larger ABR amplitudes and improved SNR (Ferm et al., 2013; Ferm et al., 2015; Wegner & Dau, 2002; Elberling et al., 2010). In this design, lower-frequency components—which require longer travel distances along the cochlea—are presented earlier than higher-frequency components. This timing strategy allows all frequency components to reach their respective cochlear regions simultaneously, resulting in improved ABR wave amplitudes and enhanced SNR, thereby reducing overall testing time (Ferm et al., 2013; Ferm et al., 2015; Wegner & Dau, 2002; Elberling et al., 2010). Consequently, the faster a stimulus achieves an adequate

Table 2. The median and inter-quartile range (IQR) for ABR acquisition time produced via NB LS CE-Chirp and tone burst stimuli at multiple frequencies and intensities using both stopping criteria in normal hearing infants.

Acquisition time (s)	Intensity (dBnHL)		Fmp			Visual	
				Frequency (Hz)			
Stimuli		500	1000	4000	500	1000	4000
NB LS CE-Chirp	70	6.01 (9.01)	4.50 (9.01)	9.01 (12.76)	12.01 (9.01)	12.01 (9.01)	9.01 (9.01)
Tone burst		16.52 (25.33)	9.01 (12.76)	6.06 (9.76)	21.02 (18.77)	21.02 (12.01)	21.02 (6.76)
NB LS CE-Chirp	40	16.52 (27.03)	12.01 (24.77)	12.01 (21.77)	15.01 (6.76)	15.01 (6.77)	10.51 (9.01)
Tone burst		33.03 (40.54)	30.03 (45.80)	15.01 (27.03)	27.03 (15.77)	30.03 (21.02)	24.02 (12.01)

Table 3. Number of ABR wave V identification using Fmp 3.1 versus number of wave V confirmed by an audiologist at frequency specifics and intensities in normal hearing infants.

Stimulus types	Stimulus frequency (Hz)	Intensity (dBnHL)	No. of wave V identified by Fmp≥ 3.1	No. of wave V confirmed by Audiologist	Percentages of discrepancies (%)
NB LS CE-Chirp	500	70	49	48	2.04
Tone burst			49	47	4.08
NB LS CE-Chirp		40	48	46	4.17
Tone burst			47	46	2.13
NB LS CE-Chirp	1000	70	50	50	0
Tone burst			50	50	0
NB LS CE-Chirp		40	50	47	4.00
Tone burst			50	45	10.00
NB LS CE-Chirp	4000	70	50	47	6.00
Tone burst			50	47	6.00
NB LS CE-Chirp		40	50	45	10.00
Tone burst			50	46	4.00

Intensity (dBnHL)	Stopping criteria	Stimulus	Frequency (Hz)	*p*-value											
				Fmp						Visual					
				NB LS CE- Chirp			Tone burst			NB LS CE- Chirp			Tone burst		
				500	1000	4000	500	1000	4000	500	1000	4000	500	1000	4000
70	Fmp	NB LS CE-Chirp	500				*<0.001			0.05					
			1000					*<0.001			*<0.001				
			4000						0.05			*0.001			
		Tone burst	500										0.138		
			1000											*0.005	
			4000												*<0.001
	Visual	NB LS CE-Chirp	500				*<0.001						*<0.001		
								0.09						*<0.001	
			4000						0.14						*<0.001
		Tone burst	500	*<0.001											
			1000		*<0.001										
			4000			*<0.001									

Table 4. The *p*-values of the post hoc analysis Wilcoxon signed-rank test for acquisition time at multiple stimulus frequencies at 70 dBnHL in normal hearing infants

Intensity (dBnHL)	Stopping criteria	Stimulus	Frequency (Hz)	*p*-value											
				Fmp						Visual					
				NB LS CE- Chirp			Tone burst			NB LS CE- Chirp			Tone burst		
				500	1000	4000	500	1000	4000	500	1000	4000	500	1000	4000
40	Fmp	NB LS CE-Chirp	500				*0.001			*<0.001					
			1000					*0.001			0.02				
			4000						0.06			*<0.001			
		Tone burst	500										0.03		
			1000											0.09	
			4000												*<0.001
	Visual	NB LS CE-Chirp	500				*<0.001						*<0.001		
			1000					*<0.001						*<0.001	
			4000						*<0.001						*<0.001
		Tone burst	500	0.030											
			1000		*<0.001										
			4000			*0.006									

Table 5. The *p*-values of the post hoc analysis Wilcoxon signed-rank test for acquisition time at multiple stimulus frequencies at 40 dBnHL in normal hearing infants

SNR, the shorter the acquisition time—regardless of the detection technique employed. Similar findings have also been reported in studies involving adult subjects (Dzulkarnain et al., 2023). The improvement of test time is important because infants are particularly vulnerable to sensory overload and extended time can cause fatigue or discomfort (Graven & Browne, 2008). The enhanced efficiency of the NB LS CE- Chirp in reducing acquisition times holds clinical significance (Cebulla & Elberling, 2010) particularly for infants, as it minimizes testing duration and discomfort. ABR amplitude is higher elicited via NB LS CE-Chirp stimulus. Therefore, the likelihood of faster ABR detection due to the reduction in the number of averaging sweeps required for effective noise attenuation. NB LS CE chirp produced higher amplitude than tone burst because the neural synchronization is more optimum and wider in NB chirp than in tone burst.

Fmp and visual detection method

The curreny study also sought comparison between two stopping criteria of ABR detection times which were visual detection and Fmp ≥ 3.1, in infants with normal hearing. The present study did not yield compelling evidence supporting the superiority of the Fmp ≥ 3.1 stopping criterion over visual detection in terms of ABR acquisition time, refer Table 3.

Furthermore, notable no discrepancies were observed between ABRs identified by the Fmp technique and those validated by an audiologist. The discrepancy is higher at lower intensity than high intensity because of poor SNR typically at low intensity levels.

To our knowledge, no published study to date has directly compared ABR detection using the Fmp method against visual detection for the infant population. These findings highlight three important considerations regarding the clinical application of the Fmp detection method in audiology practice. Objective detection techniques, such as Fmp, should be considered complementary to clinical visual interpretation in ABR assessment (NHSP, 2013; BSA, 2018). At 4000 Hz, there is no significant different between NB LS CE-Chirp and tone burst stimuli, this may be due to need more sample to identify the small differences. The expert judgment of audiologists remains the gold standard for ABR identification with Fmp values serving as useful

indicators for verifying the SNR when confirming responses. Fmp enhances the efficiency of ABR detection by allowing real-time SNR verification, thereby reducing reliance on fixed sweep counts. When Fmp used combined with visual detection, Fmp can shorten ABR testing time by concluding tests once sufficient data is collected, thus minimizing unnecessary auditory stimulation. While the Fmp cutoff value of 3.1 is appropriate at suprathreshold levels, it may not be suitable as stimulus intensity approaches threshold levels. The use of Fmp as a stopping criterion relies on SNR calculations that contribute to the robustness of the ABR waveform and facilitate faster algorithmic achievement of the Fmp threshold in NB LS CE-Chirp designs compared to tone bursts (Mckearney et al., 2010). This modified structure improves ABR waveform characteristics, leading to better testing efficiency overall. This limitation becomes more pronounced in low intensity recordings where neural synchrony and signal strength are reduced. Therefore, future studies should consider utilizing lower Fmp cutoff values corresponding to 95% or 90% confidence levels, equivalent to 2.25 and 1.80, respectively. Infants are particularly vulnerable to sensory overload and extended exposure to auditory stimuli can cause fatigue or discomfort (Graven & Browne, 2008). This approach ensures that ABR waveforms are captured during periods of optimal responsiveness of the auditory system in infants enhancing the reliability and accuracy of test results. Accurate ABR testing is crucial for early identification of hearing impairments, enabling timely intervention, and improved long term outcomes (JCIH, 2009).

CONCLUSION

This study concluded that ABR recorded from NB LS CE-chirp was faster than ABR to tone bursts at multiple frequencies and intensities, using both stopping criteria, in normal-hearing infants. The findings from the current study only focus on normal hearing infants below 6 months of age and are limited to the specific stimulus, recording parameters, and equipment used.

Future research should aim to validate these findings across broader populations, encompassing various types and severities of hearing loss. Additionally, the future study should investigate the correlation between objective ABR thresholds and behavioral hearing thresholds in older pediatric populations to assess consistency and clinical applicability.

ACKNOWLEDGEMENT

The authors wish to acknowledge the Ministry of Health Malaysia for their financial support in conducting this study. This project was partly funded by the International Islamic University Malaysia Research Matching Grant Scheme (RMGS24-004-0035). Special thanks are extended to the Department of Obstetrics and Gynaecology, SASMEC, and the Department of Otorhinolaryngology, SASMEC, for granting approval and facilitating data collection for this study.

REFERENCES

Cain, S. E., Gomes, T., Leisner, D., Lenzen., N., Rall, E., Schicke, E., and Uhler, K. M. (2020) Clinical Guidance Document Assessment of Hearing in Infants and Young Children. American Academy of Audiology Assessment of Hearing in Infant and Young Children

Cebulla, M. and Elberling, C. (2010). Auditory brain stem responses evoked by different chirps based on different delay models. Journal of American Academy Audiology. 21, 452-460.

Dzulkarnain, A. A. A., Abdullah, S. A., Ruzai, M. A. M., Ibrahim, S. H. M. N., Anuar, N. F. A., and Rahim, A. E. A. (2018). Effects of different electrode configurations on the narrow band level-specific CE-chirp and tone-burst auditory brainstem response at multiple intensity levels and frequencies in subjects with normal hearing. American Journal of Audiology, 27(3), 294–305. https://doi.org/10.1044/2018_aja-17-0087

Dzulkarnain, A. A. A., Marzuki, M. N., Shahrudin, F. A., Jamal, F. N., Chahed, N., and Zakaria, M. N. (2023). Optimization of Auditory Brainstem Response (ABR) Test Time using Level Specific (LS) CE-Chirp. International Journal of Allied Health Sciences, 7(1), 2820–2826. https://doi.org/10.31436/ijahs.v7i1.731

Dzulkarnain, A. A. A., Salamat, S., Shahrudin, F. A., Jamal, F. N., and Zakaria, M. N. (2021). Influence of Stimulus Polarity on the Auditory Brainstem Response from Level-Specific Chirp. Journal Audiology Otology, 25(2), 199-208. https://doi.org/10.7874/jao.2021.00248

Elberling, C., Callo, J., & Don, M. (2010). Evaluating auditory brainstem responses to different chirp stimuli at three levels of stimulation. Journal of the Acoustical Society of America, 128(1), 215–223. https://doi.org/10.1121/1.3397640

Ferm, I., and Lightfoot, G. (2015). Further comparisons of ABR response amplitudes, test time, and estimation of hearing threshold using frequency-specific chirp and tone pip stimuli in newborns: Findings at 0.5 and 2 kHz. International Journal of Audiology, 54(10), 745–750. https://doi.org/10.3109/14992027.2015.1058978

Ferm, I., Lightfoot, G., and Stevens, J. (2013). Comparison of ABR response amplitude, test time, and estimation of hearing threshold using frequency specific chirp and tone pip stimuli in newborns. International Journal of Audiology, 52(6), 419–423. https://doi.org/10.3109/14992027.2013.769280

Graven, S.N. and Joy V. Browne, J. V. (2008). Auditory Development in the Fetus and Infant, Newborn and Infant Nursing Reviews, 8(4), 187-193, https://doi.org/10.1053/j.nainr.2008.10.010

McKearney, R. M. and MacKinnon, R. C. (2019). Objective auditory brainstem response classification using machine learning. International Journal of Audiology, 58(4), 224–230. https://doi.org/10.1080/14992027.2018.1551633

Megha, K. N., Divyashree, K. N., Lakshmi, A., Adithya, S., Keerthana, K. P., Pushpalatha, Z. V., Konadath, S. (2019). Narrow-band chirp and tone burst auditory brainstem response as an early indicator of synaptopathy in industrial workers exposed to occupational noise. Hearing Balance Communication, 16(1): 1–12. https://doi:10.1080/21695717.2017.1418803

NHSP (2013). Sutton, G., Lightfoot, G., Contributors: Stevens, J., Booth, R., Brennan, S., Feirn, R., and Meredith, R. Guidance for Auditory Brainstem Response testing in babies.

Pinto, F.R. and Matas, C.G. (2007). A comparison between hearing and tone burst electrophysiological thresholds. Brazillian Journal Otorhinolaryngology. 73(4):513-22. https://doi: 10.1016/s1808-8694(15)30103-8

Wegner, O., and T. Dau, T. (2002). Frequency specificity of chirp-evoked auditory brain stem responses, Journal of the Acoustical Society of American. 111, 1318–1329.

Joint Committee on Infant Hearing. (2019). Year 2019 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. Journal of Early Hearing Detection and Intervention, 4(2).

ORIGINAL

The Impact of Shared Book Reading on the Speech and Language Development of Infants and Toddlers Aged 0-36 Months: A Scoping Review

Aishah Humaira' Mohd Fauzi[1], Juliana Aminah Marhaban[1] , **Mimizairul Murni Mohd Salleh[2]**

[1]Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[2]Big M Resources, 3 Jalan SG 6/9, Taman Sri Gombak, 68100 Batu Caves

ABSTRACT

Background: Shared book reading is a common activity between caregivers and their children that facilitates early speech and language development by exposing young children to rich linguistic input. While its benefits are well-documented in children over age three, limited research has explored its impact during the critical early years of 0–36 months. This scoping review aims to gather existing evidence on the effect of storybook, or shared book reading, on the speech-language development of infants and toddlers, identify effective reading strategies, and highlight gaps to inform future early intervention practices. **Methods:** A comprehensive literature search was conducted across 3 databases: ProQuest, Science Direct, and Scopus using the keywords *parental reading*, *storybook reading*, *shared-book reading*, *dialogic reading*, *speech*, *language*, *infant*, and *toddler*. Studies that were published from 2014 until 2024 were screened according to predefined inclusion and exclusion criteria, focusing on caregiver-child reading interactions in typically developing children aged 0–36 months with reported speech and language outcomes. The selection and analysis process were guided by the PRISMA-ScR framework. **Results:** This review analysed 10 studies investigating the impact of shared book reading on children aged 0–36 months. Daily reading at 6 months was linked to better expressive and receptive vocabulary at 12 months, while child engagement was a stronger predictor of expressive language than frequency alone. Print books outperformed electronic formats in promoting verbal output. Semantically-rich illustrations enhanced expressive language, while isolated images supported receptive learning. Dialogic techniques, such as questioning, labelling, and expansion, were associated with improved language outcomes across preverbal, receptive, and expressive domains as well as vocabulary development. **Conclusion:** The findings suggest that shared book reading is an effective tool that supports early speech and language development in children aged 0–36 months. The evidence supports that early and interactive reading experiences contribute meaningfully to preverbal, receptive, and expressive language.

Keywords:
parental reading; storybook reading; shared-book reading; dialogic reading; speech; language; infant; toddler

INTRODUCTION

Speech and language development in the first three years of life represents a critical period marked by rapid brain plasticity and neural growth (Shipley & McAfee, 2021). During this time, dynamic interactions between the child and their environment shape the formation and strengthening of neural pathways essential for language acquisition (Sengpiel, 2007). Repeated exposure to rich language experiences strengthens neural connections involved in language processing. Conversely, insufficient stimulation may lead to the gradual weakening and elimination of underused neural pathways, a process known as synaptic "pruning," which can adversely affect language development (Deguire et al., 2022; Tierney & Nelson, 2009; Hensch & Bilimoria, 2012; Houston et al., 2014). Thus, early sensory experiences, such as shared book reading, are vital for establishing the foundational skills necessary for effective communication.

Although various adults can engage in storybook reading, this review focuses specifically on caregiver-child interactions, as caregivers occupy a central role in the child's immediate social environment, or microsystem, as described by Bronfenbrenner's bioecological model (Bronfenbrenner, 1977, 1994). The microsystem involves direct, reciprocal interactions, known as proximal processes, that profoundly impact development (Bronfenbrenner & Morris, 2006). As primary communication partners, caregivers scaffold their child's learning, supporting language development through sustained, meaningful engagement (Shelton, 2018). Hence, the quality of caregiver-child interaction during storybook reading is a critical determinant of early language growth (Bus et al., 1995). Shared storybook

reading is defined as an activity between parent and child where the parent reads the storybook aloud to the child and engages the child in joint attention to the book's content (Vogler-Elias, 2013). Storybook reading introduces children to a wider and more diverse vocabulary than typical everyday conversations, which often revolve around immediate and concrete contexts (Salley et al., 2002; Logan et al., 2019). Through storybooks, children are exposed to decontextualized language, words, and concepts beyond their daily experiences, promoting richer linguistic input (Grolig, 2022). Studies also show that parents use more language input during shared book reading than during other play activities, significantly contributing to children's vocabulary growth (Clemens & Kegel, 2020; Richards et al., 2017).

Dialogic reading, a method in which parents actively engage children by encouraging verbal and nonverbal participation during reading, is particularly effective in fostering language development (Whitehurst et al., 1988). This approach is based on three core principles: encouraging child participation, providing contingent feedback, and tailoring the reading to the child's language level. Research demonstrates that dialogic reading yields stronger improvements in both receptive and expressive language skills compared to traditional passive reading (Salley et al., 2022; Chang et al., 2022). Randomized controlled trials further confirm its benefits in increasing children's enjoyment, attention, and engagement with storybooks (Murray et al., 2016; Salley et al., 2022). This scoping review seeks to synthesize current evidence on the effects of storybook reading on speech and language development in infants and toddlers aged 0–36 months. It highlights the importance of early language stimulation, especially within the context of parent-child interactions and dialogic reading strategies. The review aims to identify gaps in the literature related to storybook types, reading strategies, and frequency to better understand how these factors influence early language acquisition. Ultimately, the findings intend to provide caregivers, educators, and speech-language pathologists with evidence-based guidance for incorporating shared-book reading into early intervention to support optimal language development.

METHODOLOGY

Research Questions

This scoping review is guided by the following key questions:

1. What is the impact of storybook reading on the speech and language development of infants and toddlers aged 0 to 36 months?
2. What are the reported speech and language development outcomes for infants and toddlers involved in storybook reading, according to existing studies?
3. Which types of storybooks are most effective in supporting speech and language development in children aged 0 to 36 months?
4. What reading strategies and techniques are effective in promoting speech and language development in infants and toddlers within this age range?
5. How frequent should shared-book reading be implemented to be most effective in promoting speech and language development in children aged 0 to 36 months?
6. What gaps and inconsistencies exist in the current research on storybook reading and early language development?

Eligibility Criteria

Criteria of Study Participants

We included studies involving typically developing children aged 0 to 36 months. Studies involving children with developmental language delay/disorder, autism spectrum disorder, attention-deficit hyperactivity disorder, global developmental delay, or other medical diagnoses that could impede typical speech and language development were excluded. Studies that involved shared book reading activity facilitated by professional service providers such as therapists, teachers, or day care providers as the reading partners were also excluded, as the review focused on naturalistic caregiver-child interactions. Caregivers were defined as parents or other family caregivers directly involved in the child's daily interactions.

Criteria of Study Characteristics

This review included studies published in English or Malay that focused on storybook reading interventions and reported outcomes related to speech and language development. Studies were eligible if they involved comparisons between an intervention group and a control group, or if they included repeated measures such as pre- and post-intervention assessments. A wide range of study designs were accepted, including randomised controlled trials, non-randomised controlled trials, observational studies, cohort studies, case studies, and qualitative and quantitative research. However, review articles were excluded from this scoping review.

Selection of Sources of Evidence

The selection process involved three reviewers. Reviewer

1, a final-year Speech-Language Pathology student at the International Islamic University Malaysia (IIUM), conducted the primary screening. Reviewer 2, a qualified speech-language pathologist and lecturer with a PhD in Health Sciences (Speech-Language Pathology) from IIUM, provided validation and approval for the articles that passed initial screening, including the final selection of included studies. Reviewer 3, a children's book publishing consultant and founder of Swaterbit Kids Book, served as an external advisor in the selection process in the case of disagreements between Reviewer 1 and Reviewer 2.

Data Searching

A systematic search was conducted across three databases: ProQuest (28 results), ScienceDirect (12 results), and Scopus (122 results), yielding a total of 162 articles. After removing duplicates, 154 articles remained. All articles were identified using the keywords: *'parental reading,' 'storybook reading,' 'shared-book reading,' 'dialogic reading,' 'speech,' 'language,' 'infant,'* and *'toddler.'* Only articles published between 2014 and 2024 were considered. The data search began in March 2025 and was completed in May 2025.

Data Selection

During the initial screening of titles and abstracts, 135 articles were excluded for not meeting the inclusion criteria. The remaining 26 full-text articles were retrieved for further eligibility assessment. Of these, 16 were excluded for the following reasons: 2 were ongoing studies with no available results, 3 included children older than 36 months, 7 did not report on child communication outcomes, 1 was a review article, and 3 full-text articles were not accessible within the project timeframe (Figure 1).

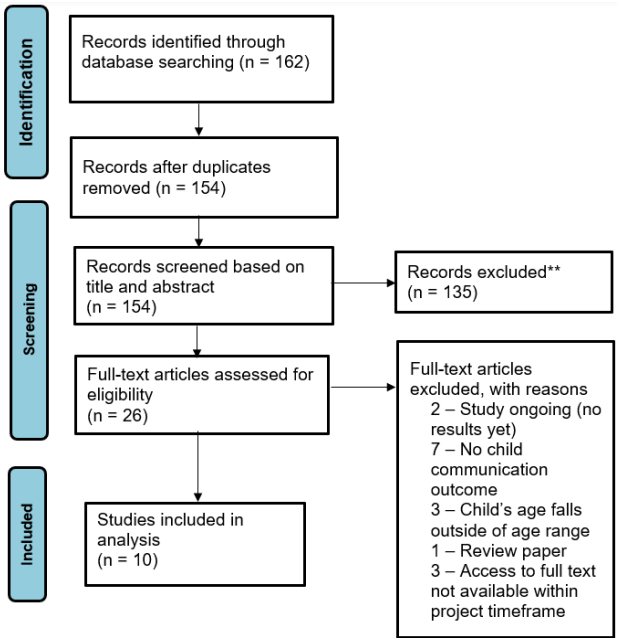


Figure 1: PRISMA flow diagram

Data Extraction/Charting Process

Data from the 10 included studies were extracted to capture key information. A thematic analysis was conducted to examine and classify the types of books used, the characteristics of the reading activities (i.e., number of books, frequency, duration, reading strategies applied), assessment methods and tools used, and any speech-language outcomes reported. To support consistency in analysis, operational definitions used for categorizing the characteristics of the storybooks are detailed in Table 1.

Data Analysis

A descriptive summary table (Table 2) was developed to present the key details of the selected studies involving shared storybook reading activities between caregivers and children, and their reported effects of storybook reading on the children's speech and language development. Table 3 summarizes key study procedures, including assessment methods and tools, intervention dosage (i.e., number of books, session frequency, and session duration), and book characteristics. Table 4 outlines the effects of shared book reading on specific language domains, as reflected in the communication

outcomes reported across studies (i.e., preverbal skills, receptive language, expressive language, vocabulary, and general communication). Table 5 details the reading strategies implemented in each study, with their definitions.

The effects were categorized as: positive findings (+), indicating improvement in communication outcomes; or ND (no difference) when no statistically significant changes were observed either pre- and post-intervention or between control and intervention groups. The notation NS (not specified) is used when no particular intervention dosage or book characteristics were described.

Table 1: Summary of book characteristics and their definitions

1.1	**Format of the book**	**Definitions**
	Basic printed picture book	Standard picture books with no interactive and manipulative features such as pop-ups, flaps, textures, and sound buttons (e.g., Just a Mess by Mercer Mayer)
	Interactive picture books	Picture books with interactive and manipulative features such as pop-ups, flaps, textures, and sound buttons (e.g., Bambu Saklambaç Oynuyor)
	Enhanced e-book	The book includes Audiovisual Hotspots—tapping an illustration reveals and narrates the word (e.g., "seagull"). It also features sound effects for images and page turns, like a dog panting or ocean waves. Holding a sentence triggers optional sentence narration.
	Basic e-book	The book features pages that can be turned by swiping, and tapping the illustrations will display the corresponding word. However, there are no sound effects and no auto-narration.
1.2	**Illustrations in the book**	**Definitions**
	Semantically rich	Illustrations that depict objects within meaningful and familiar settings (e.g, popcorn shown in a bucket at a movie theater)
	Images in isolation	Images showing objects alone, without contextual background (e.g, the same popcorn pieces displayed on a plain white background)

Table 2: Descriptive summary of included studies

ID	Study	Language (country of origin)	Study design	Child participants					Facilitator of shared book reading activity	Child's communication outcomes measured	Effect on language development
				Sample size				Age (at the time of data collection) in months	Caregiver		
				Total sample size	Intervention group	Control group	Gender		Caregiver's relationship to child		
S1	Marjanovič-Umek et al., 2016	Slovenian (Slovenia)	Longi-tudinal study	51	51	0	28M 23F	16, 31	Mother, Father	● Vocabulary size ● Mean length of utterance ● Sentence complexity (SC)	+
S2	Dicataldo & Roch, 2022	Italian (Italy)	Cross-sectional correlational study	44	44	0	28M 16F	20 - 36	Mother	◗ Receptive vocabulary ◗ Expressive vocabulary ◗ Mean length of utterance ◗ Syntax comprehension ◗ Syntax production ◗ Syntax complexity	+
S3	Bergström, 2024	Swedish (Sweden)	Feasibility study	11	11	0	6M 5F	8 - 9	Mother, Father	◗ Symbolic communication ◗ Receptive vocabulary ◗ Expressive vocabulary	+
S4	Madhavan & Mani,	German (Germany)	Quasi-experiment-	81	81	0	39F 42F	24 - 30	Mother, Father	◗ Vocabulary size ◗ Preverbal	+

	2024		al, observatio-nal and correlational study							communication: - Attention to the book - Interactiveness; frequency of shared gestures, gaze checking, and pointing › Receptive vocabulary: - Word-object mapping recognition	
S5	Ho et al., 2023	English (Australia)	Cross-secti-onal study	48	48	0	23M 26F	6, 9, and 12	Mother	› Verbal behaviours: - vocalisations › Non-verbal behaviours: - Pointing, page-turning	+
S6	Weber & Colunga, 2022	English (USA)	Longitudinal randomised control study	21	21	0	NS	19 - 23	Mother, Father	› Receptive vocabulary › Expressive vocabulary	+
S7	Dowdall et al., 2021	IsiXhosa (South Africa)	Randomised controlled trial	140	70	70	75M 65F	21 - 28	Mother, Father	› Attention › Receptive language › Expressive language › Mean length of utterance	+
S8	Munzer et al., 2018	English (USA)	Experiment-al, laboratory-based study	37	37	0	20M 17F	24 - 36	Mother, Father	› Vocabulary size › Expressive language: - Total number of verbalisations during shared-book reading	+
S9	O'Farrelly et al, 2017	English (Ireland)	Longitudinal experiment-al and quasi-	236	158	78	NS	6, 12	Mother	› Receptive vocabulary: - Total words understood	+

			experiment-al study							◗ Expressive vocabulary: - Total words produced	
S10	Ünlütabak, 2022	Turkish (Turkey)	Longitudinal study	30	30	0	12M 18F	16 - 21	Mother, Father	◗ Preverbal communication: - spontaneous gesture (pointing to a picture without parent's questions) ◗ Receptive language: - elicited gesture (pointing to pictures in response to parent's questions) ◗ Expressive language: - elicited speech (verbally responding to parent's questions) - elicited speech and gestures - spontaneous speech (speech initiations without parent's questions) - spontaneous gestures and speech (gestures and speech without parent's questions)	+

(+) indicates positive findings
(NS) indicates not specified

Table 3: Summary of study procedures

ID	Assessments conducted		Dosage of intervention			Book characteristics
	Methods	Tools	No. of books	Duration	Frequency	
S1	• 6 assessments conducted over 15 months • Vocabulary carried out as a follow-up during five-month intervals at the ages of 1;4, 1;7, 1;10, 2;1, 2;4, and 2;7 years	• Slovene adaptation of the MacArthur-Bates Communicative Development Inventories (CDI; Fenson et al.) • To assess MLU: Report from parent on the three longest utterances produced by the child	NS	NS	An average of five or six times a week	NS
S2	• Children were assessed individually in a quiet room during school hours, with an educator present.	• Test del Primo Linguaggio (TPL) - standardized language test for Italian children aged 12 and 36 months. • Italian version of the MacArthur–Bates Communicative Development Inventory (MB-CDI) • To assess MLU: Report from parent on the three longest utterances produced by the child • To assess the child's interest and engagement in reading, Principal Component Analysis (PCA) was used on three parent-reported items: Child interrupts during reading, Child invents stories during reading, Child engages in dialogue during reading	NS	NS	average at least once a day	NS

S3	◗ Assessed pre- and post-intervention ◗ First assessment conducted at 8 months, and a second assessment conducted at 18 months old	◗ Swedish version of the BAYLEY-III instrument used for assessing infant cognition, language, and motor development ◗ The communication and symbolic behavior scales development profile infant–toddler checklist (CSBS DP ITC)	NS	NS	NS	NS
S4	◗ Child's categorical interest was assessed, for each category, researchers selected: ◗ 5 familiar objects (known to be recognizable by children of the tested age) and 1 unfamiliar (novel) object to test recognition and learning ◗ Parental report and proportion of looking at objects (30 objects) were obtained to determine the child's interest. ◗ To assess the child's enthusiasm and interactiveness, video recordings of shared book-reading sessions were used. Two blind coders (new research assistants) rated the interaction quality	◗ Vocabulary size ◗ German CHILDES ◗ Wordbank database ◗ MacArthur-Bates Communicative Development Inventory (MBCDI)	6	1.3–9.96 min	NS	◗ <u>Format:</u> - E-books ◗ <u>Content Details:</u> - Images in isolation: Each object is shown against a grey background - Texts had similar length across all pages (27–31 words; mean = 28.58)

S5	• Assessment was conducted when the infants were 6, 9, or 12 months of age. • Caregivers were asked a series of questions adapted from previous studies. These included whether they had started reading with their infant, who typically participated in the reading sessions, and how frequently these sessions occurred. • The parent-child dyad participated in a 5-minute reading session in a testing room. A video recording of the interaction was obtained.	• Ages and stages questionnaire (ASQ; Squires & Bricker, 2009 - parent-report questionnaire which assesses the child's developmental progress in five sub-scales: communicati-on, gross motor, fine motor, problem-solving, and personal-social domains (Squires & Bricker, 2009)	• 11–20 books (14.6% • 21–30 books (20.8%) • 31–40 books (12.5% • 41–50 books (4.2%) • More than 50 books (47.9%)	• 1–10 min (87.5%) • 11–20 min (10.4%) • More than 30 min (2.1%)	• Almost never (2.1%) • 1–2 times per week (6.2%) • 3–4 times per week (12.5%) • 5–6 times per week (14.6%) • Every day (64.6%)	• <u>Content details:</u> - Contains 28 pages - Has a total of 230 words - Includes themes like animals, vehicles, and simple household items - Uses descriptive and rhyming storyline
S6	• Children were randomly allocated to 2 study groups: The First group received books of rich contextual illustration, the second group received books of images in isolation • Assessments were conducted at 3 time points: • Pre-intervention visit (week 1) • post-intervention visit (week 8) • follow-up visit following post-intervention (week 16)	• MacArthur-Bates Communicative Development Inventory (CDI) • List of an additional 100 food and vehicle words	4	NS	NS	• <u>Format of book:</u> - printed picture books • 2 separate books based on intervention condition, each with: Semantica-lly rich illustration • Images in isolation illustration

S7	• Assessments were conducted at 3 time points: • Baseline post-intervention (week 8) • Follow-up (6 months post-intervention)	• Early Childhood Vigilance Task (ECVT) • Short form of the MacArthur Communica-tion Developme-nt Inventory (CDI; Fenson et al., 2000) • The language subscales of the Bayley Scale of Infant Development (Bayley, 1993) • Video recording of parent-child interaction	6	10 - 15 min	everyday	NS
S8	• Researchers used a verbal coding system based on past studies of dialogic and shared e-book reading. • Analysis was done on video recordings of parent-child interactions, dividing them into 10-second intervals. Each interval was marked with a **1** if a specific verbal behavior occurred or a **0** if it did not. These scores were then totaled for each 5-minute book-reading session.	• MacArthur-Bates Communicative Developmental Inventory (CDI) short form	3	NS	NS	• <u>Format of books:</u> - Enhanced electronic book - Basic electronic books - Basic printed book
S9	• Participants are allocated into 3 groups: 1. Low-Intensity Group: Received free books as gifts starting at 3 months of age. 2. High-Intensity Group: Received	• MacArthur Communicative Development Inventories Words and Gestures (CDI; Fenson et al., 2007)	7	NS	• <u>Reading at 6 months (high intensity)</u> - Daily (41%) - A few times per week (31%) -Never/rarely (28%) • Vocabulary <u>Reading at 12</u>	NS

	books at 3 months plus home visits by trained mentors who explained the benefits of early reading, encouraged regular reading routines, boosted parents' confidence, and demonstrated and coached quality reading interactions 3. Comparison Group: Received no intervention • Assessment was conducted at 12 months of age to examine the association between shared reading at 6 months to later language skills at 12 months. • Interview with mother to obtain data on reading frequency was also conducted at 6 and 12 months of age		<u>months (high intensity)</u> - Daily (31%) - A few times per week (32%) -Never/rarely (15%) • <u>Reading at 6 months (low intensity)</u> - Daily (29%) - A few times per week (45%) -Never/rarely (46%) • <u>Reading at 12 months (low intensity)</u> - Daily (50%) - A few times per week (31%) -Never/rarely (19%)

S10	◗ The parent-toddler pairs were tested at a play center regularly attended. ◗ They were instructed to read in their usual, natural way. ◗ The book reading interactions ranged from 55s to 7 min for this task ◗ Parent-toddler conversations were transcribed word-for-word by two research assistants. Both verbal and gestural input from parents and toddlers were coded. Parent and toddler inputs were coded separately. ◗ Coding categories were based on previous research studies.	◗ Video and audio recording of parent-child interaction during shared-book reading activity	NS	NS	NS	◗ <u>Format of book:</u> - Interactive picture book with lift-the-flap feature ◗ <u>Content details:</u> - Some pages contain sentences for parents to read aloud. - Language level is slightly above the child's understanding.

(+) indicates positive findings
(NS) indicates not specified

Table 4: Summary of the effects of shared book reading by language domain

ID	Preverbal skills	Receptive language	Expressive language	Vocabulary (Re/Ex)	General communication
S1	NT	NT	+	+	NT
S2	NT	+	+	+ (Re/Ex)	NT
S3	+	+	+	NT	NT
S4	NT	+	NT	NT	NT
S5	NT	NT	NT	NT	+
S6	NT	+	+	NT	NT
S7	+	+	+	NT	NT
S8	NT	NT	+	NT	NT
S9	NT	+	+	+ (Re/Ex)	NT
S10	NT	+	+	NT	NT

(NT) indicates not tested
(+) indicates positive findings
(Re) indicates receptive
(Ex) indicates expressive

Table 5: Summary of effective reading strategies applied in each study

ID	Reading strategies applied	Definitions
S3	Contingent response	Provide immediate and appropriate responses to the child's communication attempts when the child is attentive and ready to receive them (Topping et al., 2013).
S3, S8	Expansion	Expansion means to frame and add information over and above the child's response (Whitehurst et al., 1988; Zevenbergen and Whitehurst, 2003). (e.g., Child says, "Here is a wagon," and parent replies, "a big, red wagon.")
S5, S7	Connecting	Relate the book's content to the child's personal experiences or real-world concepts.
S5, S7, S8, S10	Questions	Ask a variety of questions, such as yes/no questions, labelling questions, and open-ended questions

S7, S9	Follow child's interests/let the child take the lead	Allow the child to guide the flow of interaction by following their interest
S7, S9	Pointing and labelling	Direct the child's attention by pointing at items in the book and labeling the objects or characters.
S9	Repeat words and phrases	Expose the child to the same words and phrases repetitively to support language learning
S3	Gestures	Use non-verbal communication together with verbal communication (e.g., making a sleeping gesture while saying the word 'sleep').
S3	Onomatopoeia	Use sound-imitating words that mimic the actual sounds associated with what they represent (e.g, saying "choo-choo" instead of "train").
S3	Parentese	Use exaggerated intonation, pitch range, and elongation of sounds and vowels (Zauche et al., 2016).
S7	Exploring emotions and understanding characters' thoughts	Introduce simple emotion vocabulary (e.g., happy, angry, scared), and talk about what characters might be thinking, feeling, or planning.
S9	Encourage non-verbal communication attempts	Let the child touch, hold, and play with the books, and acknowledge or praise these attempts.
S10	Questions accompanied by pointing	Prompting the child to name objects shown in the pictures by pointing and questioning (e.g., "What's this?" while indicating with a finger).

RESULTS

This scoping review analysed ten studies that investigated the impact of shared book reading on the speech and language development in children aged 0 to 36 months. The findings are organized into five key themes: (1) assessment methods and tools used; (2) dosage of intervention; (3) book characteristics; (4) observed benefits on speech and language development; and (5) use of effective reading strategies.

Assessment Methods and Tools

All ten studies used standardized or semi-standardized tools to assess children's speech and language development (Table 3). The most commonly used assessment was the MacArthur-Bates Communicative Development Inventories (CDI), or its adapted versions, featured in seven studies (Studies S1, S2, S4, S6, S7, S8, S9). Other tools included the Test del Primo Linguaggio (TPL) (Study S2), the Swedish version of BAYLEY-III (Study S3), the Ages and Stages Questionnaire (ASQ) (Study S5), and the Communication and Symbolic Behavior Scales Developmental Profile (CSBS DP ITC) (Study S3). Researchers utilized video recordings (Studies S4, S7, S8, S10), parent-reported utterances (Studies S1, S2), observational checklists (Studies S2, S3), and coding systems (Studies S2, S4) to evaluate preverbal communication, child engagement, and the quality of parent-child interactions.

Dosage of Reading Sessions

Intervention dosage varied widely across studies, with some not explicitly stating the duration or frequency of sessions (Table 3). However, among the studies that did, reading frequency ranged from once or twice a week to daily. The most common practice reported was daily shared reading. Session durations ranged from brief interactions of 1–10 minutes to extended readings of over 20 minutes. In studies with longitudinal designs (Studies S1, S6, S9), assessments were conducted at multiple intervals to track developmental progress over time. Book exposure also varied, with one study (Study S5) reporting access to more than 50 books by nearly half of the participants.

Book Characteristics

Book types included printed picture books, electronic books (enhanced and basic), interactive books with lift-the-flap features, and themed books using rhyming storylines (Table 3). Several studies emphasized the importance of visual and thematic content. For instance, Study S6 compared books with semantically rich illustrations against those with isolated images. Other books featured common early childhood themes such as animals and vehicles. Only a few studies explicitly detailed the number of books used, but those that did reported using between 3 to 7 books, or a broader range per parental availability.

Reading Strategies

Of the studies reviewed, six described the implementation of dialogic reading strategies (Table 4), while the remainder did not detail the specific techniques used during the shared book reading interactions. Across these six studies, the most commonly reported strategies included pointing and labelling, various forms of questioning, and following the child's lead. Additional techniques encompassed expansion of child utterances, connecting the book's content to the child's personal experiences, exploring emotions, parentese, gestures, onomatopoeia, and contingent responses.

DISCUSSION

The findings addressed all the key questions by examining the overall impact of storybook reading, outlining the speech-language outcomes reported, identifying effective reading strategies and book types, and highlighting gaps within the existing literature. The evidence reinforces a growing body of research suggesting that early, interactive exposure to storybooks can positively influence multiple domains of language development, even before children begin producing full words or sentences.

What is the impact of storybook reading on the speech and language development of infants and toddlers aged 0 to 36 months?

The evidence reviewed demonstrates that shared book reading has a beneficial impact on both preverbal and verbal aspects of communication. Several studies reported gains in preverbal behaviours such as joint attention, eye gaze, vocalizations, and gesture use, which are critical precursors to expressive language (Salley et al., 2021). Most notably, improvements were also reported in receptive and expressive language, with vocabulary acquisition highlighted in three studies (S1, S2, and S9) as a key outcome. These findings reinforce that early exposure to books does more than familiarize children with print; it lays a foundational framework for linguistic, cognitive, and social development (Duursma et al., 2008).

What are the reported speech and language development outcomes for infants and toddlers involved in storybook reading, according to existing studies?

The studies reviewed collectively reported improvements across multiple language domains (Table 3). Improvements in various domains of speech and language development were reported across the included studies. Preverbal communication skills were noted to improve in two studies (S3 and S7), while growth in receptive language, including receptive vocabulary, was reported in seven studies (S2, S3, S4, S6, S7, S9, and S10). Expressive language outcomes, such as verbal output and expressive vocabulary, showed improvement in eight studies (S1, S2, S3, S6, S7, S8, S9, and S10). Additionally, three studies (S1, S2, and S9) specifically highlighted gains in vocabulary acquisition. One study (S5) reported an increase in general communication scores, although it did not specify the exact language domain assessed.

What reading strategies and techniques are effective in promoting speech and language development in infants and toddlers within this age range?

Across the studies that described reading strategies, dialogic reading techniques emerged as particularly effective. Strategies such as pointing and labelling, open-ended and WH-questions, expansion, gestures, connecting the book's content to the child's real experiences, contingent responses, and following the child's lead were frequently associated with positive language outcomes. These methods promote active engagement, reciprocal communication, and responsiveness, all of which positively impact language learning (Dong et al., 2024). Other techniques, including parentese and onomatopoeia, were also noted for promoting positive language outcomes. The consistent use of such strategies differentiates more effective shared reading interventions from passive ones (Fernald & Simon, 1984; Tamis-LeMonda et al., 2001).

How frequent should shared-book reading be implemented to be most effective in promoting speech and language development in children aged 0 to 36 months?

The reviewed studies exhibited considerable variation in the frequency and duration of shared book reading interventions. While some did not report precise dosage details, those that did generally recommended daily reading as best practice.

O'Farrelly et al. (2017) found that daily shared reading at 6 months was significantly associated with expressive and receptive vocabulary at 12 months. In contrast, reading only a few times per week was linked to gains in receptive but not expressive vocabulary, highlighting the importance of consistent, daily exposure for comprehensive language development. However, Dicataldo and Roch (2022) found no direct link between reading frequency and measures such as receptive and expressive vocabulary, syntactic comprehension, or syntactic production. Instead, they identified children's interest and engagement during reading, not frequency, as significantly related to expressive vocabulary, which in turn supports syntactic and morphosyntactic development. Frequency itself did not predict engagement, suggesting that the quality of interaction may be more influential than quantity. This is supported by Ortiz, Stowe, and Arnold (2001), who found that children whose parents were taught dialogic reading strategies showed greater engagement during shared reading than those whose parents received no such training. These findings suggest that fostering engagement through interactive techniques may have a greater impact on language outcomes than increasing reading frequency alone.

Similarly, Sočan et al. (2016) reported that shared reading frequency at 2 years and 7 months was positively associated with vocabulary size, sentence complexity, and mean length of utterance (MLU).

While daily reading remains an optimal goal, these findings underscore that its effectiveness depends on various factors, including the quality of interaction, child engagement, and caregiver input, rather than frequency alone. Future research should aim to disentangle these elements to inform reading guidelines for early intervention practices better.

Which types of storybooks are most effective in supporting speech and language development in children aged 0 to 36 months?

The types of books used across the reviewed studies varied in format, content, and interactivity, with several key characteristics emerging as influential to language outcomes. This included illustration quality, themes based on the child's interest, and book format (print versus digital).

Weber & Colunga (2022) compared semantically rich illustrated books (Table 1) to those containing isolated images and found nuanced effects on different domains of language development. Rich illustrations were associated with greater expressive language outcomes, particularly in producing target words, suggesting that visually rich contexts may facilitate children's verbal output. In contrast, books featuring isolated images were more effective for receptive vocabulary learning, as children identify and learn target words better during comprehension assessments. This highlights the importance of aligning book design with specific communicative goals—expressive versus receptive language development.

In terms of format, print books were consistently associated with more favourable outcomes than electronic books. According to Munzer et al. (2018), parents were found to use more dialogic reading strategies, such as asking questions, making storyline comments, and reading aloud more frequently during print book interactions. Conversely, during electronic book sessions (both enhanced and basic versions), these behaviors were significantly reduced (Munzer et al., 2018)

Munzer et al. (2018) also found that children demonstrated higher levels of expressive language, as measured by total verbalizations, and greater engagement during print book reading than digital formats. Books aligned with the child's interests also elicited longer attention spans and increased focus on novel word-object associations, which are critical for vocabulary learning (Madhavan & Mani, 2024). These findings suggest that content relevance and medium of presentation are key factors in optimizing the effectiveness of shared book reading experiences for children in the 0–36-month age range.

Despite these findings, few studies provided detailed specifications regarding the number or diversity of books used.

Overall, the evidence supports using developmentally appropriate, visually engaging print books with themes that align with the child's interests as a cornerstone of effective shared reading interventions. Future studies should aim to more clearly document and compare book

characteristics to refine best practices in early literacy and language development.

What gaps and inconsistencies exist in the current research on storybook reading and early language development?

While this scoping review offers valuable insights into the role of shared book reading in early speech and language development, several limitations should be acknowledged.

A notable limitation across many of the included studies was the predominance of cross-sectional designs, which restrict the ability to draw causal inferences or examine developmental trajectories over time. Longitudinal research is essential to understand how shared reading behaviours evolve and influence language outcomes over extended periods.

Several studies lacked detailed reporting on the intervention dosage, particularly about reading frequency, session duration, and strategy adherence in home settings. Although caregivers were often encouraged to read daily, few studies incorporated objective measures to verify compliance. The absence of audio or video recordings meant that data on how caregivers conducted shared reading were largely reliant on self-report, which may be prone to bias or inaccuracies.

The sample characteristics in some studies also limit generalizability. Several investigations drew participants from a single geographic location or a specific intervention programme, and a few relied on small samples, potentially underpowering their ability to detect meaningful effects. In one study, socioeconomic disparities between comparison groups may have confounded results, further complicating interpretation.

The types of books and digital reading formats varied considerably, yet were not always clearly described or standardized. These limitations highlight the need for future research to employ rigorous, longitudinal methodologies with consistent reporting of intervention fidelity and book characteristics. More robust, representative, and ecologically valid designs will strengthen the evidence base and enhance our understanding of how shared book reading can be optimized to support early speech and language development.

Implications and Future Directions

The findings of this review underscore the importance of shared book reading as a developmentally appropriate, low-cost, and accessible intervention for promoting early language skills. Clinicians, educators, and caregivers should be encouraged to adopt dialogic reading strategies during shared reading. Future research should prioritize intervention parameters (e.g., dosage, book characteristics, strategy use) and explore longitudinal effects of early shared reading on later language and literacy outcomes.

CONCLUSION

This scoping review highlights the positive impact of shared book reading on the speech and language development of children aged 0 to 36 months. The evidence supports that early and interactive reading experiences contribute meaningfully to preverbal, receptive, and expressive communication. While daily reading is optimal, child's interest and engagement, book characteristics, and caregiver strategies also play critical roles in determining outcomes. Semantically rich print books to enhance verbal output, and books with images in isolation to support receptive vocabulary, aligned with the child's interests, and used alongside dialogic techniques, are especially effective. However, variations in study design, measurement tools, and reporting practices limit broader conclusions. Future research should adopt more rigorous, longitudinal approaches to better understand how shared reading shapes early language development and to refine best practices for early intervention.

ACKNOWLEDGEMENT

Alhamdulillah. All praise be to Allah for granting me the strength, patience, and determination to complete this Final Year Project. There were moments of doubt and exhaustion, but with His guidance, I was able to see it through.

I'm truly thankful to my supervisor, Asst. Prof. Dr. Juliana binti Marhaban, and my co-supervisor, Mimizairul Murni Mohd Salleh. Thank you for your constant support, thoughtful feedback, and kind encouragement throughout this journey.

To all my lecturers and everyone at the Kulliyyah of Allied Health Sciences, thank you for the knowledge, inspiration, and opportunities you've given me over the years. Each lesson played a huge role in shaping who I am today.

To my family, thank you for being my rock. Your endless prayers, love, and support carried me through the hardest days. And to my dear friends, thank you for the laughs, the

good and bad times we spent together, and for always being there when I needed a break or a boost. I'm so grateful to have had you by my side.

This project was completed without any external funding.

From my heart, thank you to everyone who participated in this journey.

REFERENCES

Aina. (2024, July 16). Engaging Storybooks Every Preschooler Should Hear. Little Playhouse Preschool and Childcare Centre https://littleplayhouse.com.my/engaging-storybooks-every-preschooler-should-hear/

Bergström, E., Anna Sofia Bratt, & Svensson, I. (2024). An interactive shared reading intervention designed for preverbal infants: A feasibility study exploring early language and cognitive development. *Child Language Teaching and Therapy*, *40*(3). https://doi.org/10.1177/02656590241250234

Brain Plasticity & Early Intervention: "Neurons that fire together, wire together." (2022, February 10). LEARN Behavioral. https://learnbehavioral.com/blog/brain-plasticity-2

Bronfenbrenner, U. (1994). Ecological models of human development. In T. Husén & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed., Vol. 3, pp. 1643–1647). Pergamon Press.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Ed.), Handbook of child psychology: Theoretical models of human development (6th ed., Vol. 1, pp. 793–828). Wiley.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*(7), 513–531. https://doi.org/10.1037/0003-066X.32.7.513

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). *Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy*. Review of Educational Research, 65(1), 1–21. https://doi.org/10.3102/00346543065001001

Cates, C. B., Dreyer, B. P., Berkule, S. B., White, L. J., Arevalo, J. A., & Mendelsohn, A. L. (2012). Infant communication and subsequent language development in children from low-income families: The role of early cognitive stimulation. Journal of Developmental and Behavioral Pediatrics: JDBP, 33(7), 577–585. https://doi.org/10.1097/DBP.0b013e318264c10f

Chang, C. S., Hsieh, F.-J., Chen, T., Wu, S. C., Tzeng, O. J. L., & Wang, S. (2022). Revisiting dialogic reading strategies with 12-month-old infants. Early Childhood Education Journal. https://doi.org/10.1007/s10643-022-01385-4

Cheryl Jialing Ho, Duursma, E., & Herbert, J. S. (2023). Mother–infant shared book reading in the first year of life. *Infant and Child Development*. https://doi.org/10.1002/icd.2465

Deguire, F., López-Arango, G., Knoth, I. S., Côté, V., Agbogba, K., & Lippé, S. (2022). Developmental course of the repetition effect and change detection responses from infancy through childhood: a longitudinal study. Cerebral Cortex, 32(23), 5467–5477. https://doi.org/10.1093/cercor/bhac027

Dicataldo, R., & Roch, M. (2022). How Does Toddlers' Engagement in Literacy Activities Influence Their Language Abilities? *International Journal of Environmental Research and Public Health*, *19*(1), 526. https://doi.org/10.3390/ijerph19010526

Dong, Y., Chow, B. W.-Y., Mo, J., Miao, X., & Zheng, H.-Y. (2024). Effects of dialogic reading elements on children's language development. *Journal of Research in Reading*, 47(2), 181–200. https://doi.org/10.1111/1467-9817.12447

Duursma, E., Augustyn, M., & Zuckerman, B. (2008). Reading aloud to children: the evidence. *Archives of Disease in Childhood*, *93*(7), 554–557. https://doi.org/10.1136/adc.2006.106336

Dowdall, N., Murray, L., Skeen, S., Marlow, M., De Pascalis, L., Gardner, F., Tomlinson, M., & Cooper, P. J. (2021). Book-Sharing for Parenting and Child Development in South Africa: A Randomized Controlled Trial. *Child Development*, *92*(6), 2252–2267. https://doi.org/10.1111/cdev.13619

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*(1), 104–113. https://doi.org/10.1037/0012-1649.20.1.104

Karousou, A., & Economacou, D. (2024). Quantity and quality of book reading to infants and toddlers: their effect on early communication and language development. Psychology: The Journal of the Hellenic Psychological Society, 29(1), 1–27. https://doi.org/10.12681/psy_hps.34486

Lorio, C. M., Delehanty, A. D., & Romano, M. K. (2021). A Systematic Review of Parent–Child Shared Book Reading Interventions for Infants and Toddlers. Topics in Early Childhood Special Education, 42(3), 027112142199879. https://doi.org/10.1177/0271121421998793

Madhavan, R., & Mani, N. (2024). The quality of caregiver–child interaction is predicted by (caregivers' perception of) their child's interests. *Royal Society Open Science*, *11*(4). https://doi.org/10.1098/rsos.231677

Marjanovič-Umek, L., Fekonja-Peklaj, U., & Sočan, G. (2016). Early vocabulary, parental education, and the frequency of shared reading as predictors of toddler's vocabulary and grammar at age 2;7: A Slovenian longitudinal CDI study. *Journal of Child Language, 44*(2), 457–479. https://doi.org/10.1017/s0305000916000167

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. Psychological Bulletin, 137(2), 267–296. https://doi.org/10.1037/a0021890

Munzer, T. G., Miller, A. L., Weeks, H. M., Kaciroti, N., & Radesky, J. (2019). Differences in Parent-Toddler Interactions With Electronic Versus Print Books. *Pediatrics*, *143*(4). https://doi.org/10.1542/peds.2018-2012

Murray, L., De Pascalis, L., Tomlinson, M., Vally, Z., Dadomo, H., MacLachlan, B., Woodward, C., & Cooper, P. J. (2016). Randomised controlled trial of a book-sharing intervention in a deprived South African community: Effects on carer-infant interactions, and their relation to infant cognitive and

National Institute on Deafness and Other Communication Disorders. (2022). Speech and language developmental milestones. NIDCD. https://www.nidcd.nih.gov/health/speech-and-language

O'Farrelly, C., Doyle, O., Victory, G., & Palamaro-Munsell, E. (2018). Shared reading in infancy and later development: Evidence from an early intervention. *Journal of Applied Developmental Psychology*, *54*, 69–83. https://doi.org/10.1016/j.appdev.2017.12.001

Salley, B., Brady, N. C., Hoffman, L., & Fleming, K. (2019). Preverbal communication complexity in infants. *Infancy*, *25*(1), 4–21. https://doi.org/10.1111/infa.12318

Salley, B., Daniels, D., Walker, C., & Fleming, K. (2022). Shared book reading intervention for parents of infants and toddlers. Journal of Early Childhood Research, 20(3), 322–340. https://doi.org/10.1177/1476718x221091462

Sengpiel, F. (2007). The critical period. Current Biology, 17(17), R742–R743. https://doi.org/10.1016/j.cub.2007.06.017

Shelton, L. (2018). The Bronfenbrenner Primer. Routledge. Shipley, K. G., & McAfee, J. G. (2021). Assessment in speech-language pathology: A resource manual. Plural Publishing, Inc.

Tierney, A. L., & Nelson, C. A. (2009). Brain Development and the Role of Experience in the Early Years. Zero to Three, 30(2), 9. https://pmc.ncbi.nlm.nih.gov/articles/PMC3722610/

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., & Lewin, S. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Annals of Internal Medicine, 169(7), 467–473. https://doi.org/10.7326/M18-0850

Ünlütabak, B., Aktan-Erciyes, A., Yılmaz, D., Kandemir, S., & Göksun, T. (2022). Parental input during book reading and toddlers' elicited and spontaneous communicative interactions. *Journal of Applied Developmental Psychology*, *81*, 101436. https://doi.org/10.1016/j.appdev.2022.101436

Vogler-Elias, D. (2013). *Shared Storybook Reading.* 2850-2854. https://doi.org/10.1007/978-1-44191698-3_507

Weber, J., & Colunga, E. (2022). Learning from Word Books: Does the Type of Illustration Matter? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*(44). https://escholarship.org/uc/item/5kv5g42x

ORIGINAL ARTICLE

Anxiety Level Among Parents of Newly Diagnosed Children with Speech-Language Impairment: A Study Using the Malay Version of DASS.

Juliana Aminah Marhaban [1*], Nuratirah Nazihah Ngadimin[1], Rosnani Sarkarsi [2]

[1]Department of Audiology and Speech- Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[2]Department of Special Care Nursing, Kulliyyah of Nursing, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia

Keywords:
Parental anxiety; speech-language impairment; newly diagnosed children; DASS-M

ABSTRACT

Background: Parents of newly diagnosed children with speech-language impairment often experience elevated levels of anxiety and stress, primarily due to concerns about their child's developmental future and the complexity of intervention pathways. This study aims to assess the anxiety levels among these parents using the validated Malay version of the Depression Anxiety Stress Scale (DASS-M). **Methods:** A cross-sectional quantitative study was conducted involving 15 parents whose children were diagnosed with speech-language impairment within the past 12 months. Participants completed demographic questions and the DASS-M via an online Google Form. Data were analysed to identify anxiety levels and explore associations with demographic variables: educational level, socioeconomic status, and parent age group. **Results:** Most participants (86.7%) reported no anxiety, while 13.3% reported mild anxiety. Correlation analysis showed no statistically significant associations between anxiety and socioeconomic status, parent age group, and educational level. Power analysis indicated that all tests were underpowered (power < 0.80), limiting the ability to detect true relationships. **Conclusion:** This study examined anxiety levels among parents of children newly diagnosed with speech-language impairment using the DASS-M, finding that most reported no symptoms while a few exhibited mild anxiety. Despite limitations such as a small sample size and potential underreporting, the findings highlight the need for culturally sensitive tools and support strategies to better understand and address parental emotional responses during the early post-diagnosis stages.

INTRODUCTION

Speech and language are fundamental components of human communication. Speech refers to the verbal production of language, while language involves the conceptual processing of communication, including both receptive (understanding) and expressive (output) components (McLaughlin, 2011). Speech-language impairment refers to both speech-language delays and disorders (*Communication Difficulties in Childhood*, 2017).

Children with speech delays often struggle with articulation, phonological processing, or motor planning, leading to unclear or limited verbal output (American Speech-Language-Hearing Association [ASHA], 2023). On the other hand, language delays affect the child's ability to comprehend and use language appropriately for their developmental stage. These delays typically follow a normal developmental sequence but at a slower pace (Jin, 2024). On the contrary, speech or language disorders involve atypical developmental patterns and often require long-term, specialized interventions (Jin, 2024).

The prevalence of speech and language impairment is notable. Wooles et al. (2017) reported that 6-7% of children entering school present with communication impairment. Similarly, Law et al. (2000) estimated that 5-8% of preschool-aged children are affected. Globally, the prevalence may reach as high as 10-15% (Yew et al., 2014). These impairments can impede social interaction, academic achievement, and emotional development, along with communication ability (Carson et al., 1998; St Clair et al., 2011). Early identification and intervention are crucial. Early involvement by speech-language

* *Corresponding author.*
E-mail address: julia_am@iium.edu.my

Journal homepage: https://journals.iium.edu.my/ijahs/index.php/IJAHS
EISSN NO 2600-8491

pathologists and implementation of evidence-based interventions have been shown to have a significant enhancement in outcomes (Children's Hospital of Philadelphia, 2024; Law et al., 2000; Northwestern University, 2024).

The term 'newly diagnosed' is often applied to cases identified within the previous 6 to 12 months, as this time frame is frequently used in research to capture the early adjustment period (Lu et al., 2018; Dolgin et al., 2007). For parents of children with speech-language impairment, this early period following diagnosis is often marked by emotional disturbance, confusion, and heightened anxiety as they seek to understand their child's needs and navigate complex intervention pathways (Asghar et al., 2023).

Despite the significance of the early post-diagnosis period, Malaysia-based literature focusing specifically on parents' experiences during this stage for children with speech-language impairment or developmental language disorder remains sparse. Most parent-experience research in the Malaysian context emphasizes autism spectrum disorder, particularly exploring parental challenges, stigma, and support barriers (Yaacob et al., 2021). Only limited studies address speech-language impairment and parental roles in early post-diagnosis intervention (Fadzillah & Lee, 2018; *Wong et al.*, 2025), none of which foreground parents' early post-diagnosis perspectives for speec—language impairments.

Parental anxiety in response to a child's speech-language impairment is a well-documented psychological concern. According to Asghar et al. (2023), the emotional burden placed on parents stems from concerns about long-term developmental, academic, and social implications for their child. Naniwadekar (2022) further emphasizes that this anxiety is often driven by the uncertainty surrounding intervention efficacy and future communication abilities. Brannan et al. (2017) observed that parents may experience fear, frustration, and helplessness, particularly when cultural stigma or limited-service access is involved. These feelings are not only detrimental to the parent's mental health but can also affect the child's development. Fields et al. (2020) found that children of anxious parents are more likely to exhibit behavioural issues and struggle with emotional regulation due to increased household stress.

The Depression Anxiety Stress Scales (DASS; Lovibond & Lovibond, 1995) is a widely used self-report tool for measuring psychological distress. It includes subscales for depression, anxiety, and stress and has been validated across multiple cultures and populations, including in the Malay language (Hassan et al., 2020; Lovibond & Lovibond, 1995). The Malay version (DASS-M; Nurulhuda et al., 2017) has demonstrated reliability and validity within Malaysia and is particularly suited for psychological assessments involving adults. Although DASS-M has been used in psychiatric and general adult populations such as medical students (Ramli et al., 2012), nurses (Majid et al., 2019), and mixed ethnic adult patients (Musa & Aidid, 2020), its application to parents of children with speech-language impairment remains limited (Tan & Mohamad, 2019). This signals a need for more focused research to determine how well the DASS-M captures the anxiety profiles of this specific group. Despite the increasing prevalence of speech-language impairment in Malaysia (Yew et al., 2014), limited research has explored parental anxiety during the early adjustment phase following a child's diagnosis. This study aims to fill that gap by examining anxiety levels among parents of newly diagnosed children using the DASS-M.

The primary aim of this research is to determine the anxiety level among parents of newly diagnosed children with speech-language impairment. By doing so, the findings can contribute to the development of targeted support strategies that help parents manage their emotional well-being while effectively addressing their children's developmental needs.

MATERIALS AND METHODS

This cross-sectional study was conducted at the Speech and Hearing Clinic (HSC), IIUM Kuantan, and was open to participants both from the clinic and the wider community. The research spanned 4 weeks. A total of 37 parents were identified and invited to participate in the study. Of these, 21 parents provided informed consent. However, only 15 participants met the inclusion criteria and were included in the final data analysis. Six participants were excluded because they did not meet the eligibility criteria, including having a child diagnosed more than 12 months prior or having received prior mental health intervention for anxiety. The study focused on parents of children aged below 12 years who had received a diagnosis of speech-language impairment within the past 12 months. These inclusion criteria ensured the sample captured families are still in the early stages of emotional and psychological adjustment. Parents of children who had received the diagnosis more than 12 months prior, or those who had already received formal mental health interventions for anxiety, were excluded from the study. At the point of data collection, all children involved in the study had been diagnosed with speech-language impairment, with no confirmed underlying neurological, sensory, or physical conditions that could independently account for their communication difficulties.

Parents were informed about the purpose of the study and were required to provide informed consent before accessing the questionnaire. For data collection, the DASS-M questionnaire was distributed to parents via an online Google Form, which also included sections for demographic information such as the parents' age, education level, and socioeconomic status. In cases where clarification was needed due to misunderstandings or incomplete responses, the researcher followed up with the participants via email to ensure the accuracy and completeness of the data. All collected data were kept anonymous and stored securely. The responses were downloaded and analysed using the Statistical Package for the Social Sciences (SPSS).

Data Analysis

Data were analysed using the SPSS. Descriptive statistics, including frequencies and percentages were used to summarise participants' demographic characteristics and anxiety severity levels based on the DASS-M classifications. Shapiro- Wilk test was used to assess the normality of the data. As the data were not normally distributed, non-parametric statistical tests were applied. Spearman's rank-order correlation was used to examine associations between anxiety levels and selected demographic variables, including parental age group, educational level, and socioeconomic status. Statistical significance was set at $p < 0.05$. A post-hoc power analysis was conducted to evaluate the adequacy of the sample size in detecting statistically significant correlations, with a conventional power threshold of 0.80 used as a reference.

Ethical Considerations

This study received ethical approval from the International Islamic University Malaysia Research Ethics Committee (IREC) (Approval No.: IREC 2025-KAHS/DASLP14). Participation was voluntary, and informed consent was obtained from all participants prior to data collection. No personal identifiers were collected, and all data were kept confidential and used solely for research purpose

RESULTS

Participants' Characteristics and Response Rate

A total of 37 invitations were distributed to eligible parents from the HSC database. Of these, 21 parents provided consent to participate. However, only 15 participants met the inclusion criteria and were included in the final analysis.

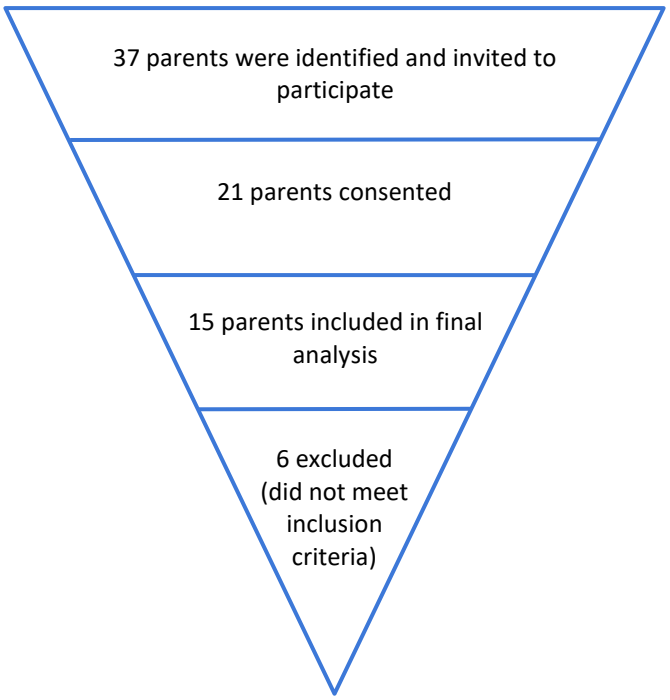


Figure 1: Participant recruitment funnel

Anxiety Level Among Parents

Descriptive analysis was conducted to summarise the frequency and percentage of participants within each severity category, providing a clear overview of the anxiety levels among the parents. Based on DASS-M, most participants (n = 13, 86.7%) reported no anxiety (Table 1). A small number of parents (n = 2, 13.3%) reported mild anxiety. No cases of moderate, severe, or extremely severe anxiety were recorded based on DASS-M severity rating.

Table 1: Participant characteristics, response rate, and anxiety levels (n= 15)

Variable	Category	No anxiety n (%) (n= 13)	Mild anxiety n (%) (n= 2)
Response rate	Invited	37	-
	Consented	21	-
	Included in final analysis	15	-
Educational level	Secondary school	4 (30.8)	0 (0.0)
	Diploma	1 (7.6)	0 (0.0)
	Degree	6 (46.2)	1 (50.0)
	Master's/ PhD	2 (15.4)	1 (50.0)

Socioeconomic status	Very low	2 (15.4)	0 (0.0)
	Low	3 (23.1)	1 (50.0)
	Lower- middle	4 (30.8)	0 (0.0)
	Upper-middle	1 (7.6)	1 (50.0)
	High	3 (23.1)	0 (0.0)
Parent age group (years)	25- 34	5 (38.5)	2 (100.0)
	35- 44	7 (53.8)	0 (0.0)
	45- 54	1 (7.6)	0 (0.0)

In terms of educational level (Table 1), among the 13 parents who reported no anxiety, nearly half held a Degree/*Ijazah Sarjana Muda* (n = 6, 46.2%). This was followed by Secondary School/*Sekolah Menengah* (n = 4, 30.8%), Masters/PhD (n = 2, 15.4%), and Diploma (n = 1, 7.6 %).

Among the 2 parents who reported mild anxiety, one (50.0%) had a Degree/*Ijazah Sarjana Muda*, and the other (50.0%) held a Master's/PhD qualification.

For the socioeconomic status (SES) classification based on Malaysian household income groups followed the Department of Statistics Malaysia (DOSM, 2020) categorization, among the 13 parents who reported no anxiety, the majority were from the Lower-Middle SES group (n = 4, 30.8%), followed by Low (n = 3, 23.1%) and High (n = 3, 23.1%). Two parents (15.4%) were from the Very Low SES group, and one parent (7.6%) was from the Upper-Middle SES group (Table 1).
Among the 2 parents who reported mild anxiety, one parent (50.0%) belonged to the Low SES group, and the other (50.0%) was from the Upper-Middle SES group.

In terms of parents' age group, based on categories used by the Department of Statistics Malaysia (DOSM, 2020), among the 13 parents who reported no anxiety, the majority were between 35-44 years old (n = 7, 53.8%), followed by those aged 25-34 years (n = 5, 38.5%) and 45-54 years (n = 1, 7.6%) (Table 4).

Among the 2 parents who reported mild anxiety, both were in the 25-34 years age group (n = 2, 100.0%).

Power of the study

A post-hoc power analysis was conducted to assess the adequacy of the sample size in detecting significant correlations among the demographic variables. The correlation between socioeconomic status and parent age group yielded a power of 0.484, indicating that the test was underpowered. Similarly, the correlation between educational level and socioeconomic status had a power of 0.228, and the correlation between educational level and parent age group showed a power of 0.084. All values were below the conventional threshold of 0.80, suggesting that the study was underpowered.

DISCUSSION

The findings of this study revealed that most parents (86.7%) reported no anxiety based on the DASS-M, while only a small proportion (13.3%) exhibited mild anxiety symptoms. This suggests that most parents in this sample may have adapted to their child's diagnosis of speech-language impairment without experiencing significant psychological distress (Burnley et al., 2023). While the DASS-M is a validated and commonly used tool, it may not capture the full extent of psychological burden experienced by parents, particularly those who are in the initial period of adjustment. Emotions such as guilt, helplessness, or fear for the future may not align precisely with the general anxiety items in the scale. Therefore, while the DASS can serve as an initial screening tool, it may not capture the full extent of psychological burden experienced by parents, particularly those who are in the initial period of adjustment to a diagnosis (Lee & Loomba, 2022).

Furthermore, this study did not account for external factors that may have influenced parental anxiety levels that were not accounted for. For instance, the severity of the child's impairment, the presence of co-occurring conditions, or whether the child had already shown progress in therapy could significantly affect parental emotional responses. Additionally, factors such as coping strategies, family or community support, and the quality of information or counselling received during the diagnostic process may have played a role. These factors may explain the low anxiety levels in some parents, not necessarily because of true emotional resilience, but due to external support or lack of awareness.

The study also raises the possibility that some parents may not have fully understood the long-term implications of the diagnosis. For parents still early in the adjustment period, lower anxiety scores may reflect limited awareness or even denial rather than emotional stability. As suggested by previous research, early engagement and education by professionals are essential to prepare parents and reduce distress (Marshall et al., 2007).

Another factor in interpreting the results is the rate of non-

participation. Although 37 invitations were distributed, only 21 parents agreed and consented to take part, and only 15 participants subsequently met the final inclusion criteria. The low response rate may indicate several potential barriers to participation. A study by Yeh et al. (2005) stated that stigma surrounding mental health, time constraints, emotional distress, and a perceived lack of relevance may discourage parents from participating in research on psychological well-being. In the context of this study, it is also possible that some parents were not yet ready to confront or reflect on their emotional reactions to their child's diagnosis, particularly during the early stage of acceptance. Future studies need to consider ways to optimize recruitment, such as offering emotional support, convenient data collection procedures, or incorporating brief psychoeducation to explain the study's relevance.

There are several strengths to this study. It focuses on a specific and often under-researched population, which is parents of newly diagnosed children with speech-language impairment, and it uses a validated tool in the Malay language, making the results more applicable to the local context. Furthermore, the focus on the early diagnosis period provides valuable insight into how parents cope during a particularly critical time. There are, however, some limitations to consider. The small sample size may restrict the generalizability of the results to other groups. Furthermore, the use of a self-report instrument like the DASS-M may have led to response bias, particularly in sociocultural environments in which discussion of mental health is still stigmatized. A further limitation is that the research did not collect in-depth data on parents' awareness and understanding of the diagnosis, or the level of information provided to them, both of which could have affected their reported anxiety scores. In addition, the study did not examine other external factors that may affect parental anxiety, such as the severity of the child's impairment, the presence of co-occurring conditions, access to therapy, or the availability of family and social support. Future research should explore the emotional experiences of these parents using a mixed-method approach. Qualitative interviews could provide deeper insights into the specific types of anxiety and concerns that parents experience following a diagnosis. It may also be useful to develop or adapt assessment tools specifically for parents of children with developmental conditions, considering cultural sensitivity and the nature of parental anxiety in this context.

CONCLUSION

This study investigated the anxiety levels among parents of children newly diagnosed with speech-language impairment using the DASS-M. Most parents in this sample reported no symptoms of anxiety, while a small proportion exhibited mild anxiety. These findings suggest that most parents had begun to emotionally adjust to their child's diagnosis during the early post-diagnostic period, though other unmeasured factors may have influenced their reported anxiety levels.

These findings suggest that while many parents may demonstrate early emotional adjustment following a chils's diagnosis, a subset may still experience psychological vulnerability. Regular screening of parental anxiety levels within the first stages of the diagnosis could help to identify those parents who would be in need of prompt emotional support. This could help to improve parental wellness, which would subsequently influence children's active involvement in speech-language therapy services.

ACKNOWLEDGEMENT

The author would like to express sincere appreciation to the staff of the Speech and Hearing Clinic, International Islamic University Malaysia (IIUM), Kuantan, for their assistance during the data collection process, and to all parents who participated in this research for their cooperation and commitment.

Gratitude is also extended to the Kulliyyah of Allied Health Sciences for providing the facilities and resources that supported the completion of this study. Finally, the author acknowledges all individuals who contributed directly or indirectly to the success of this work.

This research was not funded by any grant.

REFERENCES

American Speech-Language-Hearing Association. (2025). *American Speech-Language-Hearing Association | ASHA*. https://www.asha.org/

Asghar, S., Riaz, M., Kalsoom, U., & Mustafa, Z. (2023). Level of negative emotional states among parents of speech and language delay. *Journal of Health and Rehabilitation Research, 3*(2), 550–554. https://doi.org/10.61919/jhrr.v3i2.138

Brannan, M. S., Henker, B. R., & Nyman, M. (2017). Emotional responses of parents to the diagnosis of a speech delay in their child. *Journal of Speech, Language, and Hearing Research, 60*(4), 1052–1065.

Burnley, A., St Clair, M., Bedford, R., Wren, Y., & Dack, C. (2023). Understanding the prevalence and manifestation of anxiety and other socio-emotional and behavioural difficulties in children with developmental language disorder. *Journal of Neurodevelopmental*

Disorders, 15(1), Article 9. https://doi.org/10.1186/s11689-023-09486-w

Cambridge University Press. (2024). *Cambridge Dictionary*. https://dictionary.cambridge.org/

Children's Hospital of Philadelphia. (n.d.). *Language disorders*. Retrieved November 25, 2024, from https://www.chop.edu

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Communication difficulties in childhood. (2017). In *Taylor & Francis Knowledge Center*. https://taylorandfrancis.com/knowledge/Medicine_and_healthcare/Neurology/Speech_and_language_impairment/

Department of Statistics Malaysia. (2020). *Household income and basic amenities survey report 2019*. https://www.dosm.gov.my

Dolgin, M. J., Phipps, S., Fairclough, D. L., Sahler, O. J. Z., Askins, M., Noll, R. B., Butler, R. W., Varni, J. W., & Katz, E. R. (2007). Trajectories of adjustment in mothers of children with newly diagnosed cancer: A natural history investigation. *Journal of Pediatric Psychology, 32*(7), 771–782. https://doi.org/10.1093/jpepsy/jsm013

Fadzillah, A. J., & Lee, J. A. C. (2018). Early intervention of preschoolers with speech and language impairments by parents as first teachers: A case study on four families in Kuching, Sarawak. *Journal of Cognitive Sciences and Human Development, 3*(1), 1–10. https://doi.org/10.33736/jcshd.641.2017

Fields, M. M., Kingston, S. E., & Coles, M. E. (2020). Parental anxiety and household stress: Implications for children's emotional and behavioral regulation. *Developmental Psychology, 56*(5), 901–915. https://doi.org/10.1037/dev0000893

Hassan, S. F., Omar, S. A., & Shamsuddin, S. (2020). Validation of the Malay version of the Depression, Anxiety, and Stress Scale (DASS) among Malaysian adults. *Malaysian Journal of Medical Sciences, 27*(3), 83–92.

Jin, J. (2024). Screening for speech and language problems in young children. *JAMA, 331*(4), 368. https://doi.org/10.1001/jama.2023.27630

Law, J., Boyle, F., & Harris, A. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language & Communication Disorders, 35*(2), 165–188. https://doi.org/10.1080/136828200247133

Lee, B. K., & Loomba, R. S. (2022). Rates of depression, anxiety, and stress in parents of children with congenital heart disease using the Depression Anxiety Stress Scale. *Annals of Pediatric Cardiology, 15*(4), 374–379. https://doi.org/10.4103/apc.apc_27_22

Lu, H.-F., Chang, N.-T., Sheng, W.-H., Liao, S.-C., Wu, P.-Y., & Hsiao, F.-H. (2018). Exploring suicide ideation and suicide attempts among HIV-positive patients during the first six months to one year after their diagnosis. *The Journal of Nursing, 65*(4), 60–72. https://doi.org/10.6224/jn.201808_65(4).09

Marshall, J., Goldbart, J., & Phillips, J. (2007). Parents' and speech and language therapists' explanatory models of language development, language delay and intervention. *International Journal of Language & Communication Disorders, 42*(5), 533–555. https://doi.org/10.1080/13682820601053753

Musa, R., & Aidid, E. M. (2020). Psychometric properties of the Depression Anxiety Stress Scale 21-item (DASS-21) Malay version among a big sample population of non-Malays in Malaysia. *Malaysian Journal of Psychiatry, 29*(1), 34–42.

Naniwadekar, K. (2022). Stress and anxiety among parents of children with communication disorders. In *Research anthology on physical and intellectual disabilities in an inclusive society* (pp. 1439–1454). IGI Global. https://doi.org/10.4018/978-1-6684-3542-7.ch076

Northwestern University Center for Audiology, Speech, Language, and Learning. (2024). *Language delays and disorders*. Retrieved November 25, 2024, from https://nucasll.northwestern.edu

Nurulhuda, M. M., Ghazali, S. R., & Rahman, N. A. (2017). Validation of the Malay version of the DASS. *Journal of Clinical Psychology, 72*(2), 105–112.

Ramli, M., Rosnani, S., & Aidil Faszrul, A. R. (2012). Psychometric profile of Malaysian version of the Depressive, Anxiety and Stress Scale 42-item (DASS-42). *Malaysian Journal of Psychiatry, 21*(1), 1–9.

St Clair, M. C., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2011). A longitudinal study of behavioral, emotional and social difficulties in individuals with a history of specific language impairment (SLI). *Journal of Communication Disorders, 44*(2), 186–199. https://doi.org/10.1016/j.jcomdis.2010.09.004

Tan, E. H. F., & Mohamad, Z. S. (2019). Early intervention services for special needs children: An exploration of the effectiveness of early special education in Malaysia. *Psychological Research and Intervention, 2*(1), 11–20. https://doi.org/10.21831/pri.v2i1.24526

Wong, Y. Y., Yap, J. S., Chu, S. Y., Smith, G., Woi, P. J., Chai, S. C., Ng, L. S., & Lin, L.-Y., et al. (2025). "Parents are the First Tutors of Their Child": Parents' perception of responsibility on practicing speech-language home program during COVID-19 lockdown. *Journal Name*, *Volume* (Issue), page-range. *(in press)*

Yaacob, W. N. W., Yaacob, L. H., Zulkifli, M. M., & Muhamad, R. (2021). Behind the scenes of parents nurturing a child with autism: A qualitative study in Malaysia. *International Journal of Environmental*

Research and Public Health, 18(16), 8532. https://doi.org/10.3390/ijerph18168532

Yeh, M., McCabe, K., Hough, R. L., Hazen, A. L., & Garland, A. F. (2003). Racial/ethnic differences in parental endorsement of barriers to mental health services for youth. *Mental Health Services Research, 5*(2), 65–77. https://doi.org/10.1023/A:1023286210205

Yew, S. G., Rahmat, N. H., & Tarmizi, M. (2014). Prevalence of speech and language delay in young children in Malaysia. *Pediatrics & Child Health, 19*(5), 246–251.

REVIEW ARTICLE

A Scoping Review of Clinical Practice Guidelines for the Assessment of Sarcopenia

Siti Salwana Kamsan[1,2,*], Aliza Haslinda Hamirudin[3], Nor Syuhaidah Udin[4]

[1]Department of Physical Rehabilitation Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[2]Children Health and Wellbeing Research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[3]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia
[4]Department of Rehabilitation, Sultan Ahmad Shah Medical Center, International Islamic University Malaysia, 25200, Kuantan, Pahang, Malaysia

Keywords:
Sarcopenia; clinical practice guideline; screening; assessment

ABSTRACT

Background: Sarcopenia refers to the progressive loss of muscle mass. The adverse health effects of sarcopenia include body weakness, falls, and dependency. To reduce the complications of sarcopenia, comprehensive assessments are required for early detection and accurate diagnosis. Hence, this review aimed to synthesise clinical practice guidelines (CPGs) for sarcopenia and determine the key recommendations on instruments used for screening and diagnostic assessment. **Methods:** An extensive search was performed through the scientific electronic databases, namely Scopus, ScienceDirect, Cochrane and ProQuest Medical and Health. In addition, a grey literature search was conducted through Google Scholar to optimise the search and reduce the risk of missing CPGs. The literature search included a time frame ranging from 2016 to 2025. The CPGs were selected based on predetermined criteria. Due to the nature of the review, the findings were reported in narrative format. The protocol of this review was officially filed with the Open Science Framework. **Results:** Out of 1291 records, 8 records were included in the review for further analysis. The identified CPGs were from multiple countries, involving participation of health experts across various disciplines. Overall, the CPGs are consistent in emphasising the utilisation of SARC-F as a screening tool and assessments of muscle mass, strength, and physical function for the diagnosis and severity of sarcopenia. **Conclusion:** Screening questions can be used as a first step for screening sarcopenia and should be complemented with objective-based measures for accurate detection. The findings of this review allow clinicians and researchers to streamline the process of screening and diagnosing sarcopenia.

INTRODUCTION

Sarcopenia, also known as muscle atrophy or thinning of muscle tissue, is commonly occurs in the older population. It is a skeletal muscle syndrome, characterised by loss of muscle mass and strength (Sayer & Cruz-Jentoft 2022). The loss in muscle mass is mainly associated with the ageing process. As the age increases, the efficiency of metabolic pathways involved in protein synthesis in the muscle tissue declines, which leads to a reduction in muscle fibres and motor neurons (Cannataro et al., 2021). The progressive loss of muscle fibres may result in functional decline, independence and poor quality of life (Larsson et al., 2019).

Approximately 15% of older adults worldwide are affected by sarcopenia, and the prevalence is much higher, up to 70%, among those with illnesses (Yuan & Larsson 2023). With the global trend of an ageing population, its incidence may escalate significantly. Several regions with a rapidly ageing population have paid attention to the prevention and management of sarcopenia (Baek et al., 2023). Hence, systematic evaluation of sarcopenia is essential to ensure accurate detection and appropriate intervention.

In order to ensure accurate diagnosis of sarcopenia, a holistic assessment involving various health experts such as geriatricians, physicians, nurses, physiotherapists and nutritionists is required (Jia et al., 2025). Predominantly, the approaches of sarcopenia assessment can be broadly

categorised into objective-based measurements, which are based on performance tests, and subjective-based measurements, which do not require direct quantification (Guttikonda & Smith 2021; Ackermans et al., 2022; Tagliafico et al., 2022; Liu et al., 2024). Regardless of the nature of assessments, the selection of the instruments must be informed by evidence-based clinical standards and adhere to the clinical practice guidelines (CPGs).

Several CPGs have been established for sarcopenia. Amongst the most notable guidelines are the International CPGs for Sarcopenia (ICFSR), the Asian Working Group for Sarcopenia (AWGS), and the European Working Group on Sarcopenia in Older People (EWGSOP) (Dent et al., 2018; Liu et al., 2024). The CPGs provide a set of recommendations to standardise sarcopenia assessment and management. While there are several CPGs available in the literature, clinicians have difficulties in reviewing the guidelines due to time constraints, demanding workloads, inadequate resources, and the complex nature of the working environment (Wang et al., 2023; Verstraeten et al., 2024). To facilitate easier reference, there is a need for a review paper that provides a side-by-side comparison or a concise summary of the key recommendations on the sarcopenia assessments.

To date, there is no integrative review that summarises CPGs for sarcopenia assessment. Such a study is needed to compare, synthesise, and clarify any differences across assessment guidelines for easier clinical application. In addition, it helps to consolidate the most recent recommendations and evidence, ensuring clinical practice remains current. Hence, this scoping review aimed to collate CPGs of sarcopenia and synthesise the instruments used in the sarcopenia assessment.

MATERIALS AND METHODS

Study design

This scoping review was registered with the Open Science Framework Registry (Registration DOI: 10.17605/OSF.IO/FW5CY). The protocol of the review was guided by the Joanna Briggs Institute (JBI) guideline and Arksey and O'Malley framework (2005). To ensure transparency in reporting, the Preferred Reporting Items for Systematic Reviews and Meta-analyses extension for scoping reviews was employed (Tricco et al. 2018). Table 1 summarises the five steps involved in the protocol.

Table 1: Scoping Review Protocol suggested by Arksey and O'Malley framework

Steps	Components
1	Identifying the research question
2	Identifying relevant studies
3	Study selection
4	Charting the data
5	Collating, summarizing and reporting the results

Formulation of Research Question for the Review

The research question was formulated based on the 'population', 'concept', 'context' (PCC) framework (Peters et al. 2022). In this review, the 'population' was defined as the individuals with sarcopenia, the 'concept' is the assessment and screening instruments recommended by clinical practice guidelines, while the 'context' was the types, advantages and limitations. Hence, the review addressed two questions, which were 'What are the assessment tools used for screening sarcopenia as recommended by the clinical practice guidelines?' and 'What are the advantages and limitations of the instrument tools.

Identifying Relevant Studies (Reports)

A systematic electronic search was conducted in October 2025 through four databases, namely Scopus, ScienceDirect, Cochrane and ProQuest Medical and Health. In addition, grey literature was employed through the Google Scholar database to optimise the search and to reduce the risk of any missing reports. The search strings were developed to identify the clinical practice guidelines for sarcopenia, using three main terms, which were 'Sarcopenia', 'Clinical guideline', 'Assessment' and 'Screening. The search was performed by one researcher. The initial search results were imported, compiled and preliminarily screened through their title and abstract. Any irrelevant and duplicate studies were removed. The screening results were cross-checked and verified by another researcher.

Selection of Studies

This review includes only reports on clinical practice guidelines on sarcopenia. The reports were selected if they fulfilled the predetermined criteria, which were reports prepared by medical and health representatives, published in English, contained screening and/or assessment tools for sarcopenia and were published from 2016 onwards to ensure the relevance of the review

findings. The exclusion criteria were reports without full-text accessibility. Reports that fulfilled the criteria were exported to Excel for further analysis.

Charting and Reporting the Data

Data were systematically extracted from the included reports and were organised in a customised table. The data were categorised into metadata and key outcomes for synthesis. The recorded metadata were the characteristics of the clinical practice guidelines, including the name of the clinical practice guidelines, year of publication, country of origin, total number of screening and assessment instruments described in the reports, while the key outcomes were the names of the screening and assessment instruments, types of instruments, advantages of the instruments and limitations of the instruments. Finally, the data were narratively reported based on specific diagnosis criteria for sarcopenia as outlined in the sarcopenia diagnostic standards, set by international benchmarks.

RESULTS

Search Results

The electronic search yielded 1291 records, and 1280 records were excluded upon title and abstract screening. A total of 11 records were retrieved for full-text screening, and 8 records were eligible and were fully reviewed for final analysis. The flowchart for the selection of the records is shown in the PRISMA diagram below (Figure 1). The identified clinical practice guidelines (CPGs) were developed across multiple countries with participation from a wide range of experts, as outlined in Table 2. Across the CPGs, 7 of 8 CPGs recommended screening tests, and every guideline highlighted diagnostic assessment methods for sarcopenia detection.

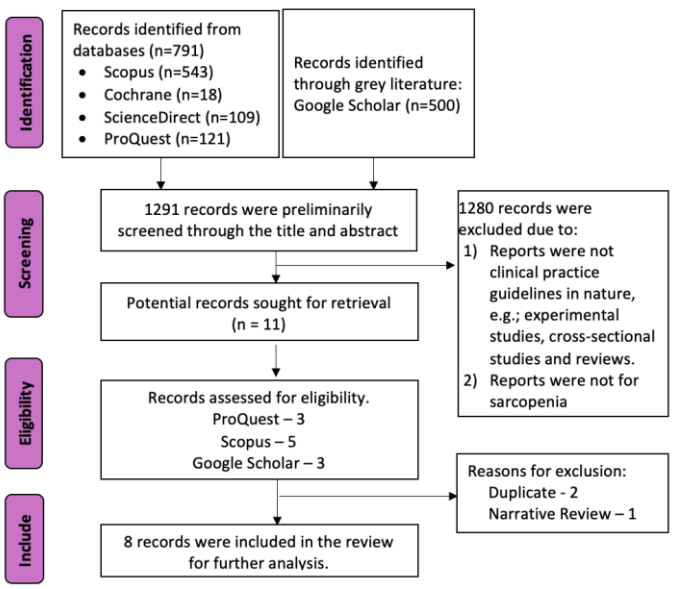


Figure 1: Prisma flow diagram.

Screening and assessment instruments for sarcopenia

The instruments were categorised into several domains, which were screening questionnaire, body composition assessment, muscle strength assessment and physical performance tests. Each of the instruments is described below, and the characteristics of the tools are outlined in Table 3.

This review found that the SARC-F questionnaire emerged as the screening tool that is most widely recommended by 7 CPGs (Kalra et al., 2025; Baek et al., 2023; Zanker et al., 2023; Lim et al., 2022; Chen et al., 2019; Cruz-Jentoft et al., 2019 & Dent et al., 2018). It is a simple instrument developed to identify individuals at risk of sarcopenia. It assesses five aspects, which are strength, walking assistance, rising from a chair, climbing stairs, and a history of falls. Another available screening tool for sarcopenia is the Ishii screening tool. This tool, however, is recommended by only one CPG (Cruz-Jentoft et al. 2019), probably due to the nature of the tool is more technical than SARC-F, as the instrument requires hand and arm circumference measurements.

Regarding body composition analysis, 7 CPGs in this review recommended bioelectrical impedance analysis (BIA) (Kalra et al. 2025, Baek et al., 2023; Zanker et al., 2023; Daly et al., 2022; Lim et al., 2022; Chen et al., 2019 & Cruz-Jentoft et al., 2019), while 8 CPGs endorsed dual-energy X-ray absorptiometry (DEXA) for sarcopenia detection (Kalra et al., 2025; Baek et al., 2023; Zanker et al., 2023; Daly et al., 2022; Lim et al., 2022; Chen et al., 2019; Cruz-Jentoft et al., 2019 & Dent et al. 2018). BIA offers a rapid, portable, and non-invasive estimation of muscle mass, while DEXA provides more precise and detailed measurements of

muscle mass, as it is an imaging-based approach. Due to the high accuracy of DEXA, it has been considered as the gold standard for body composition analysis. Nevertheless, BIA is more practical for large-scale screening or routine monitoring, particularly for community settings. Both methods, however, are essential for confirming sarcopenia after initial screening.

Calf circumference emerged as the only anthropometric measurement for sarcopenia assessment and was highlighted in 7 CPGs (Kalra et al., 2025; Baek et al., 2023; Zanker et al., 2023; Daly et al., 2022; Lim et al., 2022; Chen et al., 2019 & Cruz-Jentoft et al., 2019). It is used as a simple marker to assess leg muscle mass. This assessment approach is quick, inexpensive, and non-invasive. It helps to identify individuals at risk of sarcopenia, especially when sophisticated equipment such as DEXA or BIA is not available.

In terms of muscle strength for sarcopenia assessment, handgrip strength and sit-to-stand were recommended by all CPGs (Kalra et al., 2025; Baek et al., 2023; Zanker et al., 2023; Daly et al., 2022; Lim et al., 2022; Chen et al., 2019; Cruz-Jentoft et al., 2019 & Dent et al., 2018). Handgrip strength is used to assess upper limb strength, while the chair stand test assesses the strength and endurance of the leg muscles. Both approaches are simple, quick, and can be performed in the clinic or community, complementing muscle mass measurements for sarcopenia diagnosis. Moreover, these approaches are highly practical to be used to monitor sarcopenia progression, evaluate the effectiveness of interventions, and guide tailored exercise or nutritional programs.

The assessment of physical performance was also highlighted in all CPGs (Kalra et al., 2025; Baek et al., 2023; Zanker et al., 2023; Daly et al., 2022: Lim et al., 2022; Chen et al., 2019; Cruz-Jentoft et al., 2019 & Dent et al., 2018). Physical performance reflects how muscle loss affects mobility, balance, and daily functioning. Walking speed, the time up and go test, and the Short Physical Performance Battery were the three types of physical performance tests recommended by the CPGs. Walking speed is a simple indicator of mobility and fall risk, while the TUG test evaluates balance, agility, and lower limb strength. The SPPB, which combines balance, gait speed, and chair stand tests, offers a comprehensive assessment of overall physical function.

Table 2: Clinical Practice Guidelines for Sarcopenia

Authors	Clinical Practice Guidelines	Region	Methods	Expert Panels	Recommended Assessment
Kalra et al. (2025)	An Indian Consensus on Sarcopenia: Epidemiology, Etiology, Clinical Impact, Screening, and Therapeutic Approaches	India	Expert consensus	Experts with expertise in endocrinology, general medicine, pulmonology, intensive care, geriatric medicine, dietetics, oncology, and sports medicine	For screening (n = 1): • SARC-F questionnaire For assessment and diagnosis (n = 9): • DEXA, BIA, CT, MRI • Handheld dynamometer • Calf Circumference • Ultrasound • Gait speed • 5-chair stand test
Baek et al. (2023)	Korean Working Group on Sarcopenia Guideline: Expert Consensus on Sarcopenia Screening and Diagnosis by the Korean Society of Sarcopenia, the Korean Society for Bone and Mineral Research, and the Korean Geriatrics Society	Korea	Two rounds of Delphi process	Geriatricians, endocrinologists, physical medicine, rehabilitation specialist, oncologist, orthopaedic surgeon, family medicine, exercise physiologist, nutritionist, healthcare policy experts, and healthcare industry experts	For screening (n = 6): • SARC-F • Calf circumference • Chair stand (5-time or 30-second) • Handgrip strength • Gait speed (4 meter or 6 meter) • Time up and go test For assessment and diagnosis (n= 8): • DEXA • BIA • Handgrip strength • Short Physical Performance Battery • Gait speed (4 meter or 6 meter), • Time up and go test • Chair stand (5-time or 30-second) • 400 meter walk test
Zanker et al. 2023	Consensus guidelines for sarcopenia prevention, diagnosis and management in Australia and New Zealand	Australia and New Zealand	Four-phase modified Delphi process	Dietitian, physiotherapist, geriatrician, nurse, chiropractor, endocrinologist, exercise physiologist, geriatrician/rheumatologist, hepatologist/gastroenterologist, occupational therapist, physician,	For screening (n = 1): • SARC-F questionnaire For assessment and diagnosis (n = 10): • Handgrip strength test • Chair sit to stand test • DEXA • BIA

				psychiatrist, rehabilitation	• CT • MRI • Usual gait speed • Timed-up-and-go test • Short Physical Performance Battery • 400 meter walk test
Daly et al. (2022)	Screening, Diagnosis and Management of Sarcopenia and Frailty in Hospitalized Older Adults: Recommendations from the Australian and New Zealand Society for Sarcopenia and Frailty Research (ANZSSFR) Expert Working Group	Australia and New Zealand	Expert consensus	Geriatrics, gerontology, dietetics, exercise physiology, occupational therapy, and intensive care	For screening (n = 0): None For assessment and diagnosis (n = 14): • DEXA • BIA • CT • MRI • Ultrasound • Calf circumference • Hand dynamometer • Chair stand tests • Usual gait speed • Time up and go test • Short Physical Performance Battery • 6-minutes walk test • 400 meter walk test • Stair climb test
Lim et al. (2022)	Singapore Clinical Practice Guidelines For Sarcopenia: Screening, Diagnosis, Management and Prevention	Singapore	Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach, involving two rounds of the modified Delphi process	Geriatricians, non-geriatrician specialists, family physicians, physiotherapists, dieticians, representation from the three local healthcare clusters, namely National Healthcare Group (NHG), National University Health System (NUHS) and Singapore Health Services (SingHealth)	For case-finding (n = 5): • SARC-F questionnaire • Calf circumference • SARC-Calf • Handgrip strength • 5-time chair strength For assessment and diagnosis, follow AWGS, 2019 (n = 6): • DEXA • BIA • Handgrip strength • 5-time chair stand • 6 meter usual gait speed

					• Short Physical Performance Battery
Chen et al. 2019	Asian Working Group for Sarcopenia (AWGS)	Asia countries	Grading of Recommendations 66 Assessment, Development and Evaluation (GRADE) approach	Members of Asian Working Group for Sarcopenia – not specify the field of expertise	For screening (n = 3): • Calf circumference • SARC-F questionnaire • SARC-Calf For assessment and diagnosis (n = 6): • Handgrip strength • 5 times chair stand test • 6-metre walk • Short Physical Performance Battery • DEXA • BIA
Cruz-Jentoft et al. (2019)	Sarcopenia: revised European consensus on definition and diagnosis	European countries	Expert consensus	Members of European Geriatric Medicine Society (EuGMS) – not specify the field of expertise	For screening tools (n = 2): • SARC-F questionnaire • Ishii screening tool For assessment and diagnosis (n = 9): • Grip strength test • DEXA • BIA • CT • MRI • Gait speed • Short Physical Performance Battery • Timed-up-and-go test • 400-meter walk
Dent et al. (2018)	International Clinical Practice Guidelines For Sarcopenia (ICFSR): Screening, Diagnosis And Management	Not specified	Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach	Geriatricians, gerontologists, musculoskeletal physiologists, allied health professionals, and community dwelling older adults	For screening (n = 2): • SARC-F questionnaire • Gait speed For assessment and diagnosis (n = 3): • DEXA • Gait speed • Grip strength

AWGS - Asian Working Group for Sarcopenia; BIA - Bioelectrical impedance analysis; CT - Computed tomography; DEXA - Dual-energy X-ray absorptiometry; MRI - Magnetic resonance imaging; SARC-F - Strength, ambulation, rising from a chair, stair climbing and history of falling.

Table 3: Screening and Assessment Tools of Sarcopenia Recommended by Included CPGs

Tools	Equipment required	Advantages	Limitations	Cut-off point (Reference)
Questionnaire				
SARC-F questionnaire	• Paper-based administration: Pen, paper and clipboard or writing surface administration • Electronic administration: A tablet or any electronic device	• Inexpensive • Requires no special equipment or training • Not involve physical effort • Ideal for patient who has physical constraint	• Requiring cognitive engagement • Potential for response bias due to self-reported responses	≥ 4 (Chen et al. 2019)
Anthropometric measurement				
Calf circumference	Measuring tape and marker	• Inexpensive • Easy to administer • Not require specialize training • Useful in settings lacking advanced tools	• Less accurate in individuals with obesity or significant peripheral oedema • Does not directly measure the quality of muscle	M: <34cm F: <33cm (Chen et al. 2019)
Body composition analysis (muscle mass)				
Bioelectrical impedance analysis	Bioelectrical impedance analysis device	• Can provide estimates of fat mass, fat-free mass, and muscle mass • Only minimal training required for the clinicians • The device is portable and easy to operate • The procedure is painless and no discomfort to the patient	• Accuracy affected by hydration status, recent food intake, exercise, and skin temperature • The device requires periodically calibration to ensure accuracy • Limited ability to differentiate between muscle quality versus quantity	M: < 7.0 kg/m2 F: < 5.7 kg/m2 (Chen et al. 2019)
Dual-energy X-ray absorptiometry	Dual X-ray Absorptiometry device/machine	• High accuracy and precision for measuring muscle mass • Provides detailed regional and whole-body composition analysis	• Exposure to low-dose ionizing radiation and not suitable for repeated measurements • Expensive equipment and requires specialized facility	Height adjusted: M: < 7.0 kg/m2 F: < 5.4 kg/m2 (Chen et al. 2019)

		• Can differentiate between bone, fat, and lean tissue • Widely considered a gold standard for body composition assessment	• Less portable, typically available in hospitals or clinics • Requires trained personnel to operate the machine	BMI adjusted: M: < 0.789 F: < 0.512 (Studenski et al. 2014)
Muscles strength				
Handgrip strength	Handheld dynamometer	• Simple and quick to perform • Inexpensive to administer • Minimal training required for test administration • The device is portable and suitable for any settings • Good indicator of overall muscle strength and functional status	• Measures only upper limb strength, may not reflect the actual whole-body muscle strength • Results can be influenced by patient motivation and effort • Does not directly measure muscle mass • Not suitable for patients with cognitive impairment, upper limb dysfunction or inability to follow instructions	M: <28kg F: <18kg (Chen et al. 2019)
Sit to stand test	Chair and timer	• Simple and quick to perform • Inexpensive to administer • Minimal training required for administration • Reflects real-life activities and relevant for daily function • Provides objective measure for muscle endurance and power	• Performance can be influenced by balance, joint pain, or cardiovascular fitness • Less useful in severely frail or non-ambulatory patients • Does not directly measure muscle mass or quality	5 times sit to stand: > 10 (Yamada et al. 2022) 30 seconds: M: < 17 F: < 15 (Sawada et al. 2021)
Physical performance				
Gait speed	Timer, cone, tape measure and space to walk	• Simple and quick to perform • Inexpensive to administer		4-m walk test: Non-completion or ≥ 6 minutes for

		• Minimal training required for test administration • Reflects real-life activities, relevant for daily function • Provides objective measure of muscle endurance and power	• Performance can be influenced by balance, joint pain, or cardiovascular fitness • Less useful in severely frail or non-ambulatory patients • Does not directly measure muscle mass or quality	completion (Chen et al. 2019)
Time up and go test				≥ 12 seconds (Jung et al. 2022)
Short Physical Performance Battery	Timer, cone, chair, tape measure and space to walk	• Comprehensive assessment of lower extremity function • Combines balance, gait speed, and chair stand tests • Strongly predictive of disability, hospitalization, and mortality • Easy to administer in clinical and community settings • Only requires minimal equipment • Useful for tracking functional changes over time	• Requires some physical ability and may not be feasible for severely frail or disabled individuals • Takes longer to perform than single tests like handgrip strength or gait speed alone • Performance can be influenced by motivation, pain, or cognitive status • May be affected by comorbidities such as arthritis or neurological conditions • Requires training for standardized administration and scoring • Does not directly measure muscle mass or quality	≤ 9 points (Chen et al. 2019)

DISCUSSION

The screening tool for sarcopenia aims to identify individuals at high risk of developing sarcopenia. However, the most recommended SARC-F screening tool is subjective in nature and has limited sensitivity, which may cause false-negative results in detecting sarcopenia at its early stages (Piotrowicz et al. 2021). While SARC-F is recommended to be used as a first step to assess sarcopenia before further diagnostic assessment, the evaluation should be complemented with other instruments that are more objective and precise. Some studies recommended bypassing sarcopenia screening and proceeding directly to diagnostic assessment, especially among high-risk individuals (Voelker et al. 2021, Daly et al. 2022)

Anthropometric measurement refers to the measurement of the human body, used to assess body size, shape, and composition (Wang et al. 2024). In this review, calf circumference has been recommended by all CPGs. Studies have shown that a smaller calf circumference is associated with decreased muscle mass and strength, as well as an increased risk of sarcopenia (Kiss et al. 2024, Sato et al. 2024). Although this method is suitable for screening in the field or clinics with limited resources, it has limitations as it is affected by subcutaneous fat, oedema, and hydration status (Kiss et al. 2024). Therefore, calf circumference should be used in conjunction with other measures to improve the accuracy of sarcopenia diagnosis.

Bioelectrical impedance analysis (BIA) plays an important role in the assessment of sarcopenia because it estimates body composition objectively using the electrical resistance of body tissues (Sales et al. 2024). This method is used to calculate the skeletal muscle mass index, which helps identify muscle mass loss. Other than BIA, dual-energy X-ray absorptiometry (DEXA) is considered the alternative standard in the assessment of muscle mass for the diagnosis of sarcopenia. This method uses two different X-ray energies to accurately distinguish between bone, fat, and muscle tissue. DEXA allows the measurement of the total lean muscle mass in the body (Messina et al. 2020). Compared to BIA, DEXA offers greater accuracy, high reliability and is less affected by hydration status or body temperature. However, both BIA and DEXA only measure muscle quantity, not quality or function. Therefore, it needs to be combined with other assessments, such as muscle strength and physical function, for a comprehensive and accurate diagnosis of sarcopenia.

It has been reported that weakness, falls, and slow walking are the important clinical signs of sarcopenia (Cruz-Jentoft et al. 2019). Studies show that a decline in muscle strength occurs earlier than loss of muscle mass (Gustafsson & Ulfhake 2024), making it a more sensitive marker of physiological changes. In this review, grip strength and the chair stand test are widely recognised by CPGs as indicators of neuromuscular performance, while physical function, which is frequently measured either by walking speed, the time up and go test or the Short Physical Performance Battery (SPPB) test, reflects an individual's ability to maintain mobility and balance. Therefore, assessing strength and physical function is essential to determine the severity of sarcopenia, predict the risk of morbidity, and plan effective rehabilitation and exercise interventions.

To our best knowledge, it is the first review to systematically analyse and collate the content of Clinical Practice Guidelines (CPGs) related to sarcopenia screening and assessment instruments from various regions. The findings of this review provide guidance to clinical practitioners and researchers in improving approaches to screening and diagnosing sarcopenia more systematically, comprehensively and effectively. One limitation of this review is that the design of this scoping review did not aim to incorporate quantitative data, as the main objective was to organise and synthesise the existing evidence without undertaking a complex quantitative synthesis. Hence, this review did not assess the accuracy of the sarcopenia assessment methods recommended in each CPG.

CONCLUSION

This review has synthesised clinical practice guidelines (CPGs) related to sarcopenia assessment, providing a comprehensive overview of recommended approaches for sarcopenia detection. The findings show that all CPGs in this review are consistent in emphasising the utilisation of SARC-F as a screening tool and assessments of muscle mass, strength, and physical function for the diagnosis and severity of sarcopenia. This consistency facilitates standardisation of clinical practice across diverse geographical contexts. Overall, this review allows clinicians and researchers to streamline the process of screening and diagnosing sarcopenia.

ACKNOWLEDGEMENT

This research was not funded by any grant

REFERENCES

Ackermans, L. L. G. C., Rabou, J., Basrai, M., Schweinlin, A., Bischoff, S. C., Cussenot, O., Cancel-Tassin, G., Renken, R. J., Gómez, E., Sánchez-González, P., Rainoldi, A., Boccia, G., Reisinger, K. W., Ten Bosch, J. A., & Blokhuis, T. J. (2022). Screening, diagnosis and monitoring of sarcopenia: When to use which tool?. *Clinical nutrition ESPEN*, *48*, 36–44. https://doi.org/10.1016/j.clnesp.2022.01.027

Arksey, H., & O'malley, L. (2005). Scoping studies: Towards a methodological framework. *International journal of social research methodology, 8*(1), 19 – 32.

Baek, J. Y., Jung, H. W., Kim, K. M., Kim, M., Park, C. Y., Lee, K. P., Lee, S. Y., Jang, I. Y., Jeon, O. H., & Lim, J. Y. (2023). Korean Working Group on Sarcopenia Guideline: Expert Consensus on Sarcopenia Screening and Diagnosis by the Korean Society of Sarcopenia, the Korean Society for Bone and Mineral Research, and the Korean Geriatrics Society. *Annals of geriatric medicine and research*, *27*(1), 9–21. https://doi.org/10.4235/agmr.23.0009

Cannataro, R., Carbone, L., Petro, J. L., Cione, E., Vargas, S., Angulo, H., Forero, D. A., Odriozola-Martínez, A., Kreider, R. B., & Bonilla, D. A. (2021). Sarcopenia: Etiology, Nutritional Approaches, and miRNAs. *International journal of molecular sciences*, *22*(18), 9724.

Chen, L. K., Woo, J., Assantachai, P., Auyeung, T. W., Chou, M. Y., Iijima, K., Jang, H. C., Kang, L., Kim, M., Kim, S., Kojima, T., Kuzuya, M., Lee, J. S. W., Lee, S. Y., Lee, W. J., Lee, Y., Liang, C. K., Lim, J. Y., Lim, W. S., Peng, L. N., Sugimoto, M., Tanaka, T., Won, C. W., Yamada, M., Zhang, T., Akishita, M., Arai, H. (2020). Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia Diagnosis and Treatment. *Journal of the American Medical Directors Association*, *21*(3), 300–307.e2. https://doi.org/10.1016/j.jamda.2019.12.012

Cruz-Jentoft, A. J., Bahat, G., Bauer, J., Boirie, Y., Bruyère, O., Cederholm, T., Cooper, C., Landi, F., Rolland, Y., Sayer, A. A., Schneider, S. M., Sieber, C. C., Topinkova, E., Vandewoude, M., Visser, M., Zamboni, M., & Writing Group for the European Working Group on Sarcopenia in Older People 2 (EWGSOP2), and the Extended Group for EWGSOP2 (2019). Sarcopenia: revised European consensus on definition and diagnosis. *Age and ageing*, *48*(1), 16–31. https://doi.org/10.1093/ageing/afy169

Daly, R. M., Iuliano, S., Fyfe, J. J., Scott, D., Kirk, B., Thompson, M. Q., Dent, E., Fetterplace, K., Wright, O. R. L., Lynch, G. S., Zanker, J., Yu, S., Kurrle, S., Visvanathan, R., & Maier, A. B. (2022). Screening, Diagnosis and Management of Sarcopenia and Frailty in Hospitalized Older Adults: Recommendations from the Australian and New Zealand Society for Sarcopenia and Frailty Research (ANZSSFR) Expert Working Group. Journal of Nutrition, Health and Aging, 26(6), 637-651. https://doi.org/10.1007/s12603-022-1801-0

Dent, E., Morley, J. E., Cruz-Jentoft, A. J., Arai, H., Kritchevsky, S. B., Guralnik, J., Bauer, J. M., Pahor, M., Clark, B. C., Cesari, M., Ruiz, J., Sieber, C. C., Aubertin-Leheudre, M., Waters, D. L., Visvanathan, R., Landi, F., Villareal, D. T., Fielding, R., Won, C. W., Theou, O., Vellas, B. (2018). International Clinical Practice Guidelines for Sarcopenia (ICFSR): Screening, Diagnosis and Management. *The journal of nutrition, health & aging*, *22*(10), 1148–1161. https://doi.org/10.1007/s12603-018-1139-9

Gustafsson, T., & Ulfhake, B. (2024). Aging skeletal muscles: what are the mechanisms of age-related loss of strength and muscle mass, and can we impede its development and progression?. *International journal of molecular sciences*, *25*(20), 10932.

Guttikonda, D., & Smith, A. L. (2021). Sarcopenia Assessment Techniques. *Clinical liver disease*, *18*(4), 189–192. https://doi.org/10.1002/cld.1111

Jia, Q., Pan, Y., Pan, D., Xia, Y., & Wu, J. (2025). Application of Comprehensive Geriatric Assessment in Assessment, Management and Prognosis of Older Adult Patients with Sarcopenia: A Review of Research Progress and Practice. *Gerontology*.

Jung, H. W., Kim, S., Jang, I. Y., Shin, D. W., Lee, J. E., & Won, C. W. (2020). Screening value of timed up and go test for frailty and low physical performance in Korean older population: the Korean Frailty and Aging Cohort Study (KFACS). *Annals of geriatric medicine and research*, *24*(4), 259.

Kalra, S., Shaikh, I. A., Shende, S., Kapoor, N., Unnikrishnan, A. G., Sharma, O. P., Tiwaskar, M. H., Vora, A., Verma, S. K., Kantroo, V., Mehta, P., Lovesley, D., Sivakumar, N., Kukreja, B. B., Kulkarni, K., & Deora, A. (2025). An Indian Consensus on Sarcopenia: Epidemiology, Etiology, Clinical Impact, Screening, and Therapeutic Approaches. *International journal of general medicine*, *18*, 1731–1745. https://doi.org/10.2147/IJGM.S510412

Kiss, C. M., Bertschi, D., Beerli, N., Berres, M., Kressig, R. W., & Fischer, A. M. (2024). Calf circumference as a surrogate indicator for detecting low muscle mass in hospitalized geriatric patients. *Aging clinical and experimental research*, *36*(1), 25. https://doi.org/10.1007/s40520-024-02694-x

Larsson, L., Degens, H., Li, M., Salviati, L., Lee, Y. I., Thompson, W., Kirkland, J. L., & Sandri, M. (2019). Sarcopenia: Aging-Related Loss of Muscle Mass and Function. *Physiological reviews*, *99*(1), 427–511. https://doi.org/10.1152/physrev.00061.2017

Lim, W. S., Cheong, C. Y., Lim, J. P., Tan, M. M. Y., Chia, J. Q., Malik, N. A., & Tay, L. (2022). Singapore clinical practice guidelines for sarcopenia: screening, diagnosis, management and prevention. *The Journal of frailty & aging*, *11*(4), 348-369.

Liu, D., Wang, S., Liu, S., Wang, Q., Che, X., & Wu, G. (2024). Frontiers in sarcopenia: Advancements in diagnostics, molecular mechanisms, and therapeutic strategies. *Molecular aspects of medicine*, *97*, 101270.

Messina, C., Albano, D., Gitto, S., Tofanelli, L., Bazzocchi, A., Ulivieri, F. M., Guglielmi, G., & Sconfienza, L. M. (2020). Body composition with dual energy X-ray absorptiometry: from basics to new tools. *Quantitative imaging in medicine and surgery*, *10*(8), 1687–1698. https://doi.org/10.21037/qims.2020.03.02

Peters, M. D. J., Godfrey, C., McInerney, P., Khalil, H., Larsen, P., Marnie, C., Pollock, D., Tricco, A. C., & Munn, Z. (2022). Best practice guidance and reporting items for the development of scoping review protocols. *JBI evidence synthesis*, *20*(4), 953–968. https://doi.org/10.11124/JBIES-21-00242

Piotrowicz, K., Głuszewska, A., Czesak, J., Fedyk-Łukasik, M., Klimek, E., Sánchez-Rodríguez, D., Skalska, A., Gryglewska, B., Grodzicki, T., & Gąsowski, J. (2021). SARC-F as a case-finding tool for sarcopenia according to the EWGSOP2. National validation and comparison with other diagnostic standards. *Aging clinical and experimental research*, *33*(7), 1821–1829. https://doi.org/10.1007/s40520-020-01782-y

Sales, W. B., Mâcedo, S. G. G. F., Gonçalves, R. S. D. S. A., de Andrade, L. E. L., Ramalho, C. S. T., de Souza, G. F., & Maciel, Á. C. C. (2024). Use of electrical bioimpedance in the assessment of sarcopenia in the older aldults: A scoping review. *Journal of Bodywork and Movement Therapies*, *39*, 373-381.

Sato, R., Sawaya, Y., Hirose, T., Shiba, T., Yin, L., Tsuji, S., ... & Urano, T. (2024). Measurement of the calf muscle circumference is useful for diagnosing sarcopenia in older adults requiring long-term care. *Annals of Geriatric Medicine and Research*, *29*(1), 58.

Sawada, S., Ozaki, H., Natsume, T., Deng, P., Yoshihara, T., Nakagata, T., Osawa, T., Ishihara, Y., Kitada, T., Kimura, K., Sato, N., Machida, S., & Naito, H. (2021). The 30-s chair stand test can be a useful tool for screening sarcopenia in elderly Japanese participants. *BMC musculoskeletal disorders*, *22*(1), 639. https://doi.org/10.1186/s12891-021-04524-x

Sayer, A. A., & Cruz-Jentoft, A. (2022). Sarcopenia definition, diagnosis and treatment: consensus is growing. *Age and ageing*, *51*(10), afac220.

Studenski, S. A., Peters, K. W., Alley, D. E., Cawthon, P. M., McLean, R. R., Harris, T. B., Ferrucci, L., Guralnik, J. M., Fragala, M. S., Kenny, A. M., Kiel, D. P., Kritchevsky, S. B., Shardell, M. D., Dam, T. T., & Vassileva, M. T. (2014). The FNIH sarcopenia project: rationale, study description, conference recommendations, and final estimates. *The journals of gerontology. Series A, Biological sciences and medical sciences*, *69*(5), 547–558. https://doi.org/10.1093/gerona/glu010

Tagliafico, A. S., Bignotti, B., Torri, L., & Rossi, F. (2022). Sarcopenia: how to measure, when and why. *La Radiologia medica*, *127*(3), 228–237.

Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. Ann Intern Med. 2018;169:467–473

Verstraeten, L. M., van Wijngaarden, J. P., Meskers, C. G., & Maier, A. B. (2024). High sarcopenia awareness contrasts a lack of clinical implementation among geriatric rehabilitation health care professionals in the Netherlands: EMPOWER-GR. *Journal of Geriatric Physical Therapy*, *47*(2), 67-76.

Voelker, S. N., Michalopoulos, N., Maier, A. B., & Reijnierse, E. M. (2021). Reliability and concurrent validity of the SARC-F and its modified versions: a systematic review and meta-analysis. *Journal of the American Medical Directors Association*, *22*(9), 1864-1876.

Wang, T., Tan, J. B., Liu, X. L., & Zhao, I. (2023). Barriers and enablers to implementing clinical practice guidelines in

primary care: an overview of systematic reviews. *BMJ open*, *13*(1), e062158.

Wang, M., Song, Y., Zhao, X., Wang, Y., & Zhang, M. (2024). Utilizing anthropometric measurements and 3D scanning for health assessment in clinical practice. *Physical activity and health*, *8*(1), 182-196.

Yamada, M., Lim, J. Y., Assantachai, P., Tanaka, T., Kim, M., Lee, S. Y., Lim, W. S., & Arai, H. (2022). Five-repetition sit-to-stand test: End with the fifth stand or sit?. *Geriatrics & gerontology international*, *22*(4), 362–364. https://doi.org/10.1111/ggi.14358

Yuan, S., & Larsson, S. C. (2023). Epidemiology of sarcopenia: Prevalence, risk factors, and consequences. *Metabolism*, *144*, 155533.

Zanker, J., Sim, M., Anderson, K., Balogun, S., Brennan-Olsen, S. L., Dent, E., Duque, G., Girgis, C. M., Grossmann, M., Hayes, A., Henwood, T., Hirani, V., Inderjeeth, C., Iuliano, S., Keogh, J., Lewis, J. R., Lynch, G. S., Pasco, J. A., Phu, S., Reijnierse, E. M., … Scott, D. (2023). Consensus guidelines for sarcopenia prevention, diagnosis and management in Australia and New Zealand. *Journal of cachexia, sarcopenia and muscle*, *14*(1), 142–156. https://doi.org/10.1002/jcsm.13115

ORIGINAL ARTICLE

The Comparison of Contact Lens Wear Clinical Symptoms between Different Modality of Lenses

Nur Amani Izzati binti Mat Ghani[1], Noor Ezailina binti Badarudin [2,*]

[1]Department of Optometry and Visual Science, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Bandar Indera Mahkota, 25200 Pahang, Malaysia
[2]Faculty of Health and Life Science, Management and Science University, 40100 Selangor, Malaysia

Keywords:
dry eye; contact lens; soft contact lens; disposable; CLDEQ-8

ABSTRACT

Background: Dry eye is a common reason for discomfort and discontinuation among soft contact lens (SCL) wearers. This study compared the frequency and intensity of contact-lens–related symptoms between disposable and conventional SCL wearers in Kuantan, Malaysia. **Methods:** This cross-sectional study recruited thirty-two SCL wearers aged 18–37 years during routine aftercare visits at an optometry clinic. Participants were categorized into disposable (n = 16) and conventional (n = 16) SCL groups based on their optometrist-prescribed lens modality. The Women's Health Study Questionnaire (WHSQ) served as a preliminary screening, followed by the Contact Lens Dry Eye Questionnaire–8 (CLDEQ-8) to assess symptom frequency and intensity. **Results:** Mann–Whitney U tests showed no significant differences between disposable and conventional lens wearers for the frequency or intensity of dryness, discomfort or blurry vision (all p-values > 0.05). Dryness was the most frequent symptom, while discomfort showed the highest intensity. Spearman correlation analyses demonstrated strong associations between symptom frequency and intensity, particularly for dryness ($r_s = 0.882$, $p < 0.001$), followed by discomfort ($r_s = 0.449$, $p = 0.013$) and blurry vision ($r_s = 0.520$, $p = 0.003$). **Conclusion:** Dryness, discomfort and blurry vision remain prevalent among SCL wearers. The strong correlation between symptom frequency and intensity highlights the need for routine symptom monitoring during aftercare visits to improve patient comfort and compliance.

INTRODUCTION

Contact lens (CL) wear has gained increasing popularity among teenagers and young adults. Globally, an estimated 140 million people wear contact lenses (Stapleton, 2009) with approximately 6–7% of the Malaysian population using them mainly for refractive correction (Mohidin & Fung, 2009). Over the past two decades, international prescribing surveys have shown a significant shift toward daily disposable modalities due to their convenience, comfort and potential benefits for ocular health (Efron, Morgan & Woods, 2013; Morgan et al., 2024; Morgan & Nichols, 2025).

Despite these advantages, many wearers report symptoms such as dryness, discomfort and fluctuating vision which often leading to reduced wearing time or lens removal to seek relief (Begley et al., 2000; Chalmers et al., 2012). Poor compliance with lens care practices as reported among Malaysian wearers (Bhandari & Rou, 2012) may further exacerbate these complications.

Among all symptoms, dryness is the most frequently reported and is a leading cause of contact lens discontinuation (Chalmers et al., 2006; Richdale et al., 2007). The TFOS Dry Eye Workshop identified contact lens wear as a significant risk factor for dry eye, characterized by discomfort, visual disturbance and tear film instability (DEWS, 2007; TFOS DEWS II, 2017). Previous studies have also shown that dryness and discomfort worsen toward the end of the day due to tear film instability, lens dehydration, deposit buildup and reduced blinking (Begley et al., 2001; Guillon & Maissa, 2005; McMonnies, 2013). Consequently, understanding symptom patterns is crucial for effective patient management and informed lens selection.

Validated questionnaires such as the Contact Lens Dry Eye Questionnaire-8 (CLDEQ-8) and the Women's Health Study Questionnaire (WHSQ) are commonly used to assess dry eye symptoms (Nichols et al., 2002; Schaumberg et al., 2003). In this study, the WHSQ served as a preliminary screening tool, while the CLDEQ-8 assessed frequency and intensity of contact-lens-related symptoms. Despite improvements in lens design, dryness remains the most

* *Corresponding author.*
E-mail address: author.mail@gmail.com

persistent complaint among SCL users, underscoring the need for routine monitoring (Dumbleton et al., 2013; Lazon de la Jara & Efron, 2022).

Although disposable lenses are generally believed to reduce dryness compared to conventional lenses, findings remain inconsistent across studies (Wolffsohn et al., 2010). Evidence comparing these modalities in the Malaysian population is also limited. Therefore, this study aimed to compare the frequency and intensity of contact-lens-related dry eye symptoms among disposable and conventional soft lens wearers using validated questionnaires to provide clinically relevant insight into symptom patterns within this population.

MATERIALS AND METHODS

Study Design and Participants

This cross-sectional study was conducted at the Optometry Clinic of the International Islamic University Malaysia (IIUM). Thirty-two soft contact lens (SCL) wearers were recruited during routine aftercare visits. Participants were categorized into two groups based on the lens modality prescribed and verified by the optometrist: disposable (n=16) and conventional (n=16). Disposable lenses were defined as those replaced daily, bi-weekly or monthly, while conventional lenses were defined as those designed for extended use, typically up to one year.

Study Procedure and Questionnaires

After providing written informed consent, eligible participants completed a questionnaire set. The study followed four sequential stages: identification, screening, eligibility assessment and final inclusion as illustrated in Figure 1.

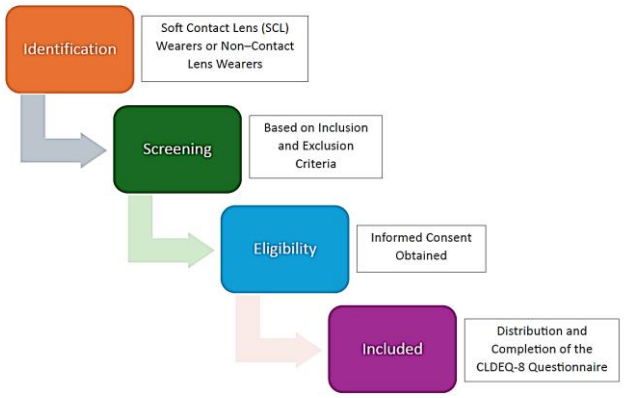


Figure 1: Flow chart of the study

Identification

Potential participants were identified from patients attending routine aftercare visits at the clinic. To be considered for screening, individuals had to be current SCL wearers aged between 18 and 40 years.

Screening

The WHSQ was used as an initial screening tool to identify individuals with symptoms suggestive of ocular dryness. Participants responded to three items addressing previous dry eye diagnosis, frequency of dryness and frequency of irritation, using categorical response options such as *Yes/No* or *constantly, often, sometimes, never*.

Although WHSQ was originally developed for large-scale epidemiological studies in women, it was employed in the present study as a gender-neutral screening tool to identify individuals with self-reported ocular dryness.

Eligibility Asessment

Participants who met the inclusion criteria based on WHSQ responses then had their medical history reviewed. Individuals with a history of ocular surface disease, systemic conditions affecting tear film stability, previous ocular surgery or those using medications known to induce dry eye symptoms were excluded at this stage.

Final Inclusion and Data Collection

Participants who passed the eligibility assessment were included in this study. Then, the Contact Lens Dry Eye Questionnaire-8 (CLDEQ-8) must be completed, which assessed the frequency and intensity of five key symptoms: eye discomfort, dryness, changeable or blurry vision, the urge to close the eyes and the need to remove lenses. Each item was scored on a Likert-type scale from 0 (*never/not intense*) to 5 (*constantly/very intense*). All questionnaires were completed under researcher supervision to address any participant questions without influencing responses.

Ethical Considerations

The study protocol was approved by the IIUM Research Ethics Committee (IREC). All procedures adhered to the principles of the Declaration of Helsinki. Participants received a thorough briefing before enrolment and provided written informed consent prior to data collection.

Figure 1: Flow chart of the study

Statistical Analysis

Data were analyzed using SPSS version 12.0.1 (SPSS Inc., Chicago, IL, USA). Descriptive statistics were presented as mean ± standard deviation. Given the small sample size and ordinal nature of the questionnaire data, non-parametric tests were used for analysis. The Mann-Whitney U test was applied to compare symptom scores (frequency and intensity) between the disposable and conventional lens groups. Spearman's rank correlation coefficient was used to assess the relationship between symptom frequency and intensity. A p-value of less than 0.05 was considered statistically significant.

RESULTS

The study included 32 participants, evenly divided between disposable (n=16) and conventional (n=16) contact lens wearers.

Comparison between disposable and conventional SCL wearers

Descriptive statistics for symptom frequency and intensity are presented in Table 1. Disposable lens wearers consistently reported lower mean scores across all symptomatic parameters compared to conventional wearers as shown in Figure 2. However, Mann-Whitney U tests revealed no statistically significant differences between the two groups for any parameter (all p-values > 0.05).

Specifically, no significant differences were observed in discomfort (U = 110.00, Z = –0.75, p = 0.515), dryness (U = 126.00, Z = –0.08, p = 0.956), blurry vision (U = 114.50, Z = –0.54, p = 0.616) or the tendency to close the eyes (U = 124.50, Z = –0.14, p = 0.897). Likewise, there were no significant differences in the frequency of removing lenses (U = 88.00, Z = –1.60, p = 0.138), intense discomfort (U = 105.00, Z = –0.90, p = 0.402), intense dryness (U = 126.00, Z = –0.08, p = 0.956), or noticeable blurry or foggy vision (U = 120.50, Z = –0.29, p = 0.780).

Table 1: Comparison of symptom frequency and intensity between disposable and conventional contact lens wearers

Symptoms	Disposable (n=16) Mean ± SD	Conventional (n=16) Mean ± SD	p-value
Frequency of discomfort	1.38 ± 0.72	1.56 ± 0.81	0.453
Frequency of dryness	1.56 ± 1.03	1.56 ± 1.26	0.938
Frequency of blurry vision	1.19 ± 0.98	1.38 ± 1.09	0.591
Frequency of closing eyes	1.19 ± 1.22	1.19 ± 0.91	0.891
Frequency of lens removal	1.88 ± 1.31	2.50 ± 1.41	0.110
Intensity of discomfort	1.94 ± 1.29	2.00 ± 1.09	0.369
Intensity of dryness	1.94 ± 1.39	2.00 ± 1.41	0.939
Intensity of blurry vision	1.31 ± 1.01	1.50 ± 1.41	0.770

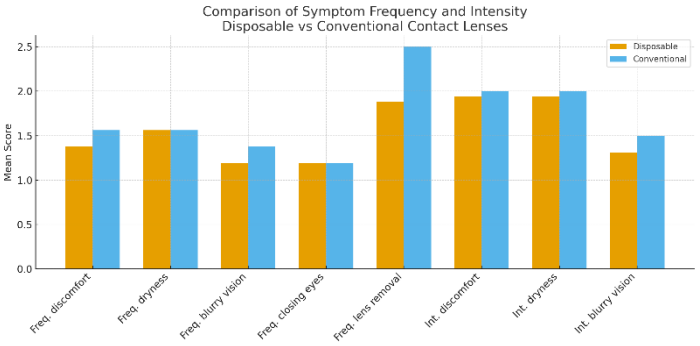


Figure 2: The comparison of symptom frequency and intensity of the disposable and conventional contact lenses

Symptom Patterns and Correlations

Dryness was the most frequently reported symptom, whereas discomfort was reported as the most intense symptom among all wearers. The disposable lens group indicated that while the frequency of symptoms was relatively low, the intensity was perceived as higher when symptoms occurred.

Spearman correlation analysis revealed significant positive relationships between the frequency and intensity of key dry eye symptoms (Figure 2). The association was strongest for dryness ($r_s = 0.882$, $p < 0.001$), followed by blurry vision ($r_s = 0.520$, $p = 0.003$) and discomfort ($r_s = 0.449$, $p = 0.013$).

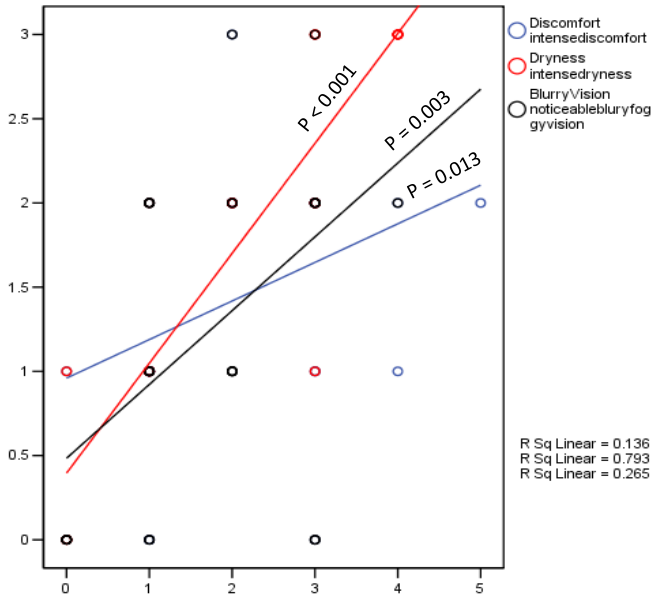


Figure 2: The association between the frequency and intensity of the dry eye symptoms

DISCUSSION

The present study investigated dry eye symptoms among disposable and conventional CL wearers using the CLDEQ-8 and the Women's Health Study Questionnaire (WHSQ). The findings revealed that dryness was the most prominent and frequently reported symptom in both groups, followed by discomfort, blurry vision and the need to close the eyes. This pattern is consistent with previous studies, which highlight dryness and discomfort as the primary complaints among SCL wearers (Fonn, 2007; Pult et al., 2009).

Although the WHS Questionnaire identified participants with a previous dry eye, most continued to wear their lenses despite reporting symptoms. This suggests that symptom severity was insufficient to cause discontinuation. Previous studies have reported similar findings, where up to 40–50% of SCL wearers experienced dryness or discomfort but did not remove their lenses (Pult et al., 2009). This behaviour may reflect adaptation strategies, symptom tolerance or the mitigating effects of advances in lens material and design (Morgan et al., 2009).

Recent evidence supports the notion that newer contact lens materials and designs especially silicone hydrogel and water-gradient lenses can reduce dryness symptoms. For example, a study evaluating Lehfilcon A silicone hydrogel water-gradient daily disposables demonstrated improved tear film stability and reduced CLDEQ-8 scores in both low- and high-symptom groups (Farias et al., 2023). Moreover, a 2024 comparative trial found that Verofilcon A lenses provided significantly better tear breakup times, lower conjunctival hyperemia and higher subjective comfort compared to Narafilcon A among wearers reporting ocular discomfort (Güell et al., 2024).

In previous studies, the prevalence of dry eye is known to increase with age particularly among women and individuals over 40 years (Albietz, 2000; Lin et al., 2003). However, the present sample consisted primarily of younger adults, which may explain the relatively low discontinuation rate despite reported symptoms. Younger wearers generally have more stable tear film function than older populations, thereby tolerating dryness better. Nonetheless, differences in diagnostic criteria may also account for variations in prevalence estimates across studies (Schein et al., 1997).

Disposable and Conventional Contact Lenses

The study found no statistically significant difference in dry eye symptoms between disposable and conventional lenses. This may be attributed to the limited sample size, which restricted the statistical power to detect differences. However, the trend indicated that conventional lens wearers reported slightly higher scores for dryness and discomfort compared to disposable CL wearers.

This observation is consistent with Marshall et al. (1992), who reported that conventional lenses are more prone to protein accumulation due to extended replacement cycles, potentially increasing ocular discomfort. In contrast, disposable or frequently replaced lenses minimize deposit formation and are associated with fewer complications (Fonn, 2007). The present study supports this notion as disposable CL wearers generally reported lower symptom intensity.

Disposable Contact Lenses

Within the disposable lens group, dryness emerged as the most prevalent and pronounced symptom, aligning with previous reports indicating that hydrogel lenses tend to lose water content during wear, thereby contributing to ocular discomfort. The introduction of silicone hydrogel lenses with higher oxygen transmissibility has enhanced comfort, particularly toward the end of the day (Morgan et al., 2009). The current findings suggest that while symptoms persist, disposable lenses especially those made of silicone hydrogel may help maintain comfort and reduce the lens removal.

Behavioural factors such as hygiene practices and compliance also play an important role in symptom perception. Wu et al. (2010) highlighted the impact of lens care behaviour on ocular health outcomes. Similarly,

previous surveys in Malaysia reported poor compliance with lens care practices among wearers (Bhandari & Rou, 2012) which could contribute to symptom persistence. In the present study, the finding that symptoms were common yet not severe enough to be reported as a reason for discontinuation. This suggests that comfort and lifestyle needs outweigh mild to moderate dryness. Sapkota et al. (2015) similarly observed that wearers often persist with lens use particularly with disposable modalities.

An additional finding was the statistically significant relationship between symptom frequency and intensity. This reinforces the understanding that greater symptom frequency is typically accompanied by higher intensity, which together contribute to the subjective burden of dry eye in SCL wearers.

The results emphasize that dryness is a critical clinical indicator in the assessment of CL-related dry eye and should be routinely monitored during aftercare. Although no statistically significant differences were found between lens modalities, the trend toward lower symptom burden in disposable wearers supports the clinical preference for frequent replacement schedules to enhance comfort and ocular surface health.

To our knowledge, this is the first study to evaluate dry eye symptoms in Malaysian SCL wearers using both the WHSQ and CLDEQ-8. By combining validated questionnaires, the study provides localized evidence that contributes to the global literature on contact lens-related dry eye and highlights the importance of symptom monitoring in clinical practice.

CONCLUSION

This study demonstrates that dryness is the most prominent dry eye symptom among both disposable and conventional SCL wearers. Although there was no significant difference between lens modalities, conventional wearers tended to report higher intensity of dryness and discomfort, consistent with previous literature linking longer replacement schedules to greater ocular surface complications. The relationship identified between symptom frequency and intensity underscores the importance of regular monitoring and patient education.

The findings support the current trend of prescribing disposable lenses as both appropriate and beneficial, supporting global shifts in practice patterns (Long et al., 1998; Morgan et al., 2009). For clinicians, raising awareness about the early signs of contact lens-related dry eye is crucial to promote patient adherence and long-term ocular health.

Future research should build upon this study by including larger and more diverse populations, and by exploring the role of environmental and behavioural factors in symptom variability. The use of standardized tools such as CLDEQ-8 and WHS questionnaires can provide valuable insights for both practitioners and researchers in understanding and managing contact lens-related dry eye.

ACKNOWLEDGEMENT

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. The author acknowledge the use of an artificial intelligence language model (ChatGPT) for language refinement during the preparation of this manuscript.

REFERENCES

Albietz, J. M. (2000). Prevalence of dry eye subtypes in clinical optometry practice. *Optometry and Vision Science*, *77*(7), 357–363.

Begley, C., Caffery B., Kinney, K., Portello, J., Davis, L., Simpson, T., & Charmers, R. (1998). Results of a dry eye questionnaire (DEQ) from optometric practices in the United States and Canada. *Optometry and Vision Science,* 75(12s), 865.

Begley, C. G., Caffery, B., Nichols, K. K., Chalmers, R. (2000). Responses of contact lens wearers to a dry eye survey. *Optometry and Vision Science, 77*(1), 40–46.

Begley, C. G., Chalmers, R. L., Mitchell, G. L., Nichols, K. K., Caffery, B., Simpson, T., Du Toit, R., Portello, J., & Davis, L. (2001). Characterization of ocular surface symptoms from optometric practices in North America. *Cornea, 20* (6), 610–618.

Bhandari, M., & Rou, H. P. (2012). Habits of contact lens wearers toward lens care in Malaysia. *Medical Journal of Malaysia*, *67*(3), 274–277.

Brennan, N. A., & Efron, N. (1989). Symptomatology of HEMA contact lens wear. *Optometry and Vision Science,* 66, 834–838.

Caffery, B., Richter, D., Simpson, T., Fonn, D., Doughty, M., & Gordon, K. (1998). CANDEES. The Canadian

Dry Eye Epidemiology Study. *Advances in Experimental Medicine and Biology,* 438, 805–806.

Cavanagh, H. D., Robertson, D. M., Petroll, W. M., & Jester, J. V. (2010). Castroviejo Lecture 2009: 40 years in search of the perfect contact lens. *Cornea*, 29, 1075–1085.

Chalmers, R. L., & Begley, C. G. (2006). Dryness symptoms among an unselected clinical population with and without contact lens wear. *Contact Lens Anterior Eye,* 29, 25-30.

Chalmers, R. L., Begley, C. G., Moody, K., & Hickson-Curran, S. B. (2012). Contact Lens Dry Eye Questionnaire-8 (CLDEQ-8) and Opinion of Contact Lens Performance. *Optometry and Vision Science*, 89(10), 1435–1442.

Craig, J. P., Nichols, K. K., Akpek, E. K., Caffery, B., Dua, H. S., Joo, C. K., ... & Wolffsohn, J. S. (2017). TFOS DEWS II Definition and Classification Report. *Ocular Surface, 15*(3), 276–283.

Doughty, M. J., Fonn, D., Richter, D., Simpson, T., Caffery, B., Gordon, K. (1997). A patient questionnaire approach to estimating the prevalence of dry eye symptoms in patients presenting to optometric practices across Canada. *Optometry and Vision Science,* 74 ,624–631.

Dumbleton, K. A., Richter, D., Woods, C. A., Aakre, B. M., Plowright, A., Morgan, P. B., & Jones, L. W. (2013). A multi-country assessment of compliance with daily disposable contact lens wear. *Contact Lens and Anterior Eye*, 36(6), 304–312.

Efron, N., Morgan, P. B., & Woods, C. A. (2013). An international survey of daily disposable contact lens prescribing. *Clinical and Experimental Optometry,* 96, 58–64.

Farias, C. C., Ogata, N. G., Matayoshi, S., & Yamane, I. S. (2023). Tear film dynamics between low and high CLDEQ-8 score with Lehfilcon A silicone hydrogel water gradient contact lens. *Diagnostics, 13*(5), 939. https://doi.org/10.3390/diagnostics13050939

Fonn, D. (2007). Targeting contact lens induced dryness and discomfort: What properties will make lenses more comfortable. *Optometry and Vision Science,* 84(4), 279–285.

Guillon, M., & Maissa, C. (2005). Dry eye symptomatology of soft contact lens wearers and nonwearers. *Optometry and Vision Science*, 82(9), 829–834. https://doi.org/10.1097/01.opx.0000177805.56766.51

Güell, J. L., Madrid-Costa, D., & Pardo, A. (2024). Comparison of comfort and tear stability with two novel silicone hydrogel daily disposable contact lenses. *Contact Lens & Anterior Eye, 47*(2), 101923. https://doi.org/10.1016/j.clae.2024.101923

Lazon de la Jara, P., & Efron, N. (2022). Contact lens comfort and compliance: Recent insights. *Contact Lens and Anterior Eye, 45*(5), 101558.

Lin, P. Y., Tsai, S. Y., Cheng, C. Y., Liu, J. H., Chou, P., & Hsu, W. M. (2003). Prevalence of dry eye among an elderly Chinese population in Taiwan: The Shihpai Eye Study. *Ophthalmology*, *110* (6), 1096–1101. https://doi.org/10.1016/S0161-6420(03)00262-8

Long, B., Brennan, N. A., & Efron, N. (1998). A global survey of contact lens prescribing, 1997. *Contact Lens and Anterior Eye, 21*(4), 73–78.

Marshall, E. C., Begley, C. G., & Nguyen, C. H. D. (1992). Frequency of complications among wearers of disposable and conventional soft contact lenses. *International Contact Lens Clinic*, 19(3-4), 55–60.

McMonnies, C. W. (2013). Psychological and other mechanisms for end-of-day soft lens symptoms. *Optometry and Vision Science,* 90(6), e175–e181.

Mohidin, N., & Fung, T. L. (2009). A Survey of Optometric Contact Lens Prescribing in Malaysia. *Jurnal Sains Kesihatan Malaysia*, 7(2), 59–72.

Morgan, P. B., Woods, C. A., Tranoudis, I., Efron, N., Santodomingo-Rubido, J., Nichols, J. J., & Jones, L. W. (2024). International trends in daily disposable contact lens prescribing (2000–2023): An update. *Contact Lens and Anterior Eye, 47*(6), 101960. https://doi.org/10.1016/j.clae.2024.101960

Morgan, P. B., & Nichols, J. J. (2025). International contact lens prescribing in 2024. *Contact Lens Spectrum, 40*(1), 26–31. Retrieved from https://www.clspectrum.com/issues/2025/januaryfebruary/international-contact-lens-prescribing-in-2024/

Morgan, P. B., Plowright, A. J., Moody, K. J., & Henderson, T. (2009). A daily disposable silicone hydrogel contact lens in neophyte wearers. *Optometry and Vision Science,* 86, E-abstract 95814.

Nichols, J. J., Mitchell, G. L., Nichols, K. K., Chalmers, R., & Begley, C. (2002). The performance of the contact lens dry eye questionnaire as a screening survey for contact lens-related dry eye. *Cornea,* 21 (5), 469–475.

Nichols, J. J., Ziegler, C., Mitchell, G. L., & Nichols, K. K. (2005). Self-reported dry eye disease across refractive modalities. *Investigative Ophthalmology & Visual Science,* 46(6), 1911– 1914.

Orsborn, G. N., & Zantos, S. G. (1989). The relationship of contact lens discomfort to symptoms of dryness. *International Contact Lens Clinic, 16*(9), 225–229.

Papas, E. B. (2014). The role of oxygen in contact lens discomfort. *Contact Lens & Anterior Eye, 37*(1), 2–9.

Pult, H., Murphy, P. J., & Purslow, C. (2009). A novel method to predict the dry eye symptoms in new contact lens wearers. *Optometry and Vision Science,* 86(9), E1042–E1050.

Richdale, K., Sinnott, L. T., Skadahl, E., & Nichols, J. J. (2007). Frequency of and factors associated with contact lens dissatisfaction and discontinuation. *Cornea, 26* (2), 168–174. https://doi.org/10.1097/ICO.0b013e31802b492d

Sapkota, K., Martin, R., Franco, S., & Lira, M. (2015). Common symptoms of Nepalese soft contact lens wearers: A pilot study. *Journal of Optometry*, 8(1), 6–11.

Schaumberg, D. A., Sullivan, D. A., Buring, J. E., & Dana, M. R. (2003). Prevalence of dry eye syndrome among US women. *American Journal of Ophthalmology,* 136, 318–326.

Schein, O. D., Munoz, B., Tielsch, J. M., Bandeen-Roche, K., & West, S. (1997). Prevalence of dry eye among the elderly. *American Journal of Ophthalmology,* 124, 723–728.

Stapleton, F., Keay, L., Jalbert, I., & Cole, N. (2007). The epidemiology of contact lens related infiltrates. *Optometry and Vision Science,* 84, 257-272.

Stapleton, F., Tan, J., Papas, E., Evans, V. E., & Keay, L. (2017). Contact lens-related complications: The effect of lens replacement frequency. *Contact Lens & Anterior Eye, 40*(3), 133–140.

The International Dry Eye Workshop. (2007). The definition and classification of dry eye disease: Report of the Definition and Classification Subcommittee of the International Dry Eye Workshop. *The Ocular Surface, 5*(2), 75–92.

Wolffsohn, J. S., Hunt, O. A., & Chowdhury, A. (2010). Objective clinical performance of "comfort-enhanced" daily disposable soft contact lenses. *Contact Lens and Anterior Eye*, 33(2), 88–92.

Wu, Y., Carnt, N., & Stapleton, F. (2010). Contact lens user profile, attitudes and level of compliance to lens care. *Contact Lens and Anterior Eye*, 33(4), 183–188.

Young, G., Chalmers, R. L., Napier, L., Hunt, C., & Kern, J. (2011). Characterizing contact lens-related dryness symptoms in a cross-section of UK soft lens wearers. *Contact Lens and Anterior Eye,* 34, 64–70.

Young, G., Chalmers, R., Napier, L., Kern, J., Hunt, C., & Dumbleton, K. (2012). Soft contact lens-related dryness with and without clinical signs. *Optometry and Vision Science*, 89(8), 1125–1132.