Translation, cross-cultural adaptation and validation of User Mobile App Rating Scale (uMARS)

Tuan Yuswana Tuan Soh[1,3], Nik Mohd Mazuan Nik Mohd Rosdy[2], Budi Aslinie Md Sabri[1]*

[1]Centre of Population Oral Health and Clinical Prevention, Faculty of Dentistry, Universiti Teknologi MARA, 47000 Sungai Buloh, Selangor, Malaysia.
[2]Centre of Oral and Maxillofacial, Diagnostics & Medicine Studies, Faculty of Dentistry, Universiti Teknologi MARA, 47000 Sungai Buloh, Selangor, Malaysia.
[3]Alor Gajah District Dental Health Office, Melaka.

Abstract

The user Version Mobile apps rating scale was established to evaluate the mobile apps. However, to date, there is no uMARS in the Malay version. This study aims to develop a Malay language alternative to the existing version of the Mobile Apps Rating Scale. The initial version of uMARS had previously underwent cross-cultural adaptation, and forward-backward translation with synthesis discussion through a development phase. The upgraded Malay version has been screened and rated by 10 respondents for face validation and a total of 36 respondents contributed to the internal reliability assessment by answering the pilot study question. All items and constructs in the uMARS version were fully adapted. All items and constructs from the prior version of uMARS were fully incorporated into the recent version. The Malay language version of uMARS was subsequently assessed for validity as well as reliability after undergoing forward backward translation. Scale level face validity index based on average method (S-FVI/Ave): 0.99, and S-FVI based on universal agreement method (S-FVI/UA): 0.89 showed that uMARS Malay Version has achieved a satisfactory level of response process validity. Whereas all items and construct presented with excellent internal reliability, Cronbach alpha (α) = 0.918, 0.857, 0.984 for objective quality, subjective quality and perceived impact. The Malay language of uMARS represents the outcome produced through proper development and validation of questionnaires; all of which favourably resulted in an updated version of uMARS that has been deemed competent to be utilized for qualitative measurement of mobile health apps in the Malay language.

Keywords: *digital dentistry, digital health records, mobile health, mobile application rating scale, reliability and validity*

Introduction

Aside from being able to readily provide health care and information to patients, the technology also plays a monumental role in bridging the supply-demand gap between patients and healthcare providers whilst simultaneously enhancing the learning process (Barbosa & Marin, 2009; Sharma *et al*., 2021). The trend of using mobile applications is expected to have an inevitable omnipresence due to the undeniably large fraction of the human population being smartphone owners who are also increasingly mindful about utilizing health-oriented applications (Subramani Parasuraman *et al*., 2017). The use of these mobile apps has benefited both parties in improving communication and service

delivery (Park *et al*., 2016; Zhang *et al*., 2018).

Their triple threat of being popular, mobile and technologically capable propagates the further development of mobile health apps to aid healthcare service providers in solidifying the continuity of care. Mobile technologies are advancing at a rapid rate, suitable for delivering health interventions at the individual level (Perez-Jover *et al*., 2019; Pretlow *et al*., 2015). Several researchers have collectively concluded technology to be a viable aid for health providers in the task of delivering health related information to patients (Heifetz & Lunsky, 2018; Perez-Jover *et al*., 2019; Shao *et al*., 2014; Soler *et al*., 2009; Underwood *et al*., 2015). However, digital product development is faced with two possibilities, success or failure (DeLone & McLean, 1992). Therefore, evaluating such information systems is very important (Delone & McLean, 2003; Mang *et al*., 2022).

Due to the lack of assessment of mobile health apps, it is unknown if the ever-increasing research in health includes the content necessary for health generally or specifically. Therefore, the need for evaluation has become a significant challenge. One way of measuring them is through a valid and reliable questionnaire (Price *et al*., 2015). uMARS is a tool to provide evaluation and measurement scales related to user feedback in evaluating application development from heuristic aspects, user experience and usability to eHealth and mHealth interventions (Stoyanov *et al*., 2016). Numerous scholars have utilized uMARS to evaluate various mobile applications, including contact tracing apps centered on smoking cessation as well as oncological applications (Bendotti *et al*., 2022; Berhanuddin *et al*., 2017; Lu *et al*., 2021; Sereda *et al*., 2019; Sharpe & Kamara, 2018; Strodl *et al*., 2020). The uMARS scale has demonstrated its user-friendliness and may be used as a legitimate and reliable indicator of the app quality. A key challenge in adapting the uMARS to Malay is ensuring that the translated items retain their original meanings. Words, phrases, and health terms often carry different meanings across languages, which could affect how users understand them. Health-related terms, in particular, need careful selection to capture the right nuances and avoid confusion that can result from direct translation (Beaton *et al*., 2000). Cultural context also shapes adaptation; Malay-speaking users may have unique expectations due to their cultural beliefs and habits around health management and information-seeking. Including culturally relevant details in the Malay uMARS can help it better reflect these preferences, leading to more accurate assessments (Sousa & Rojjanasrirat, 2011). Additionally, varied literacy levels in Malay-speaking populations can impact accessibility, so simplifying complex terms or adding clarifications helps ensure that the tool works for all education levels (Cha *et al*., 2007). Adapting the uMARS isn't just about translation, it requires testing to confirm the tool's accuracy and consistency while translating in other language such Malay (Hernández *et al*., 2020). By meeting local standards, the Malay uMARS can help build user trust and remain relevant to Malay-speaking populations (Bardus *et al*., 2020). This version is thus expected to improve health app evaluations, ensuring they meet the cultural, ethical, and usability needs of local users and ultimately support better healthcare through high-quality mobile health solutions (Stoyanov *et al*., 2016).

It is significant to note that the questionnaire showed satisfactory reliability and validity and was brief and easy to use. It's vital to acknowledge the questionnaire's satisfactory level of reliability and validity; both of which further strengthen its existing traits of being user-friendly. uMARS has a good intraclass correlation coefficient (0.66 - 0.70) over a one- to two-month period and great internal consistency = 0.90. Taking into account the risks associated with app inaccuracies as well as the potential exploitation of user data as a result of data breaches. With these factors in mind, it is believed that mobile apps for local use should be appraised using a suitable and understandable tool to gauge the accurate data for app advancement. In addition, the terminology barrier may make it challenging

to effectively capture the local target population's thoughts, feelings, perceptions, behaviours, and attitudes. Studies on the quality of the apps using validated uMARS have been published in a number of languages, including Japanese, Spanish, Italian, and Arabic (Bardus *et al.*, 2020; Martin-Payo *et al.*, 2021; Morselli *et al.*, 2021; Shinohara *et al.*, 2022). However, none of these have been developed or adapted and validated to specifically measure the quality of mobile apps in the Malay language. Therefore, this study aims to 1) develop a Malay version of the user version of the Mobile Apps Rating Scale. 2) assess the uMARS Malay version's reliability and validity in evaluating mobile apps. The valid and accurate Malay version of the uMARS questionnaire will be helpful in gauging the calibre of mobile apps among the Malay-speaking populace, particularly in nations like Malaysia, Indonesia, Brunei, and Singapore.

Materials and Methods

Study design

The adaptation and validation of the uMARS Malay Version were conducted using the approach described by (Sousa & Rojjanasrirat, 2011). The approach can be divided into 1) cross-cultural adaptation and translation process, 2) Assessment of Reliability and validity. uMARS is a measurement scale related to user feedback in evaluating application development from heuristic aspects, user experience and usability to eHealth and mHealth interventions (Stoyanov *et al.*, 2016). There are 26 items used to evaluate the objective quality, subjective quality and perceived impact. Each uMARS item uses a 5-likert scale (1-Inadequate, 2-Poor, 3-Accepted, 4-Good, 5-Excellent) except item number 13-16, there is a "not applicable" option. Overall, the objective quality of uMARS covers four different structures namely: engagement (items 1–5), functionality (items 6-9), aesthetics (items 10–12), and information (items 13–16). Scores are derived respectively from each item to obtain the average score of each structure (dimension). The total MARS score is also calculated as the mean score for the first four dimensions. The subjective quality score was obtained as the mean score for items 17-22, and perceived impact has six questions in item F and focuses on health impact.

Cross-cultural adaptation and translation process

It was decided that all 26 items and 6 constructs were adapted and maintained for translation. The uMARS structure is illustrated in Figure 1.
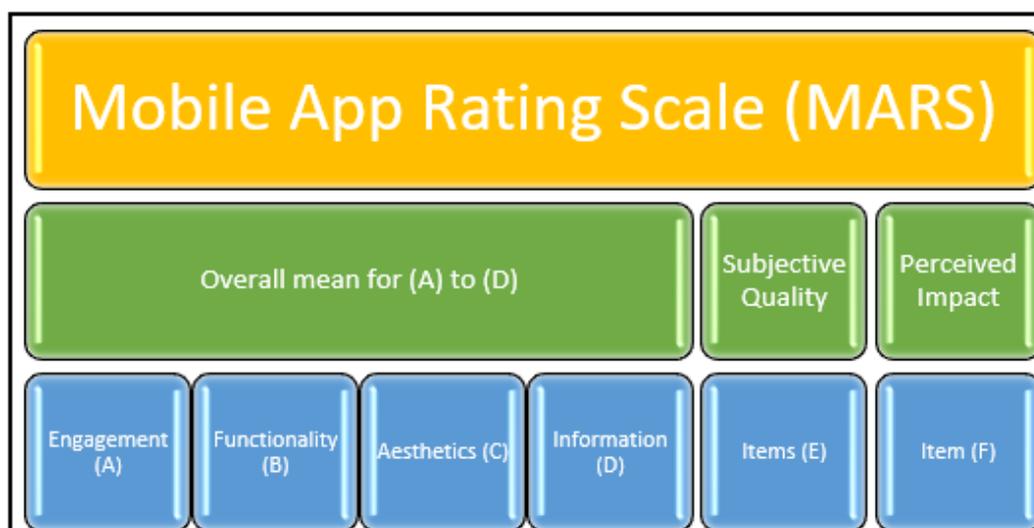

Figure 1. Items in Mobile Apps Rating Scale (MARS).

The translation process for the uMARS involved selected bilingual experts and cognitive interview participants to ensure accuracy and cultural relevance. Bilingual experts included a linguistic specialist and healthcare professionals familiar with mobile health terminology, chosen for their proficiency in both English and Malay, along with experience in translation or cross-cultural adaptation. For cognitive interviews, participants were selected from a broad range of potential users, including healthcare providers, patients, students, and lecturers, to reflect the diversity of the Malay-speaking audience. This approach ensured that the adapted uMARS would be understandable, relevant, and applicable across different user groups.

The adapted uMARS was translate from English to the Malay language by three independent translators, of which one of them is a linguistic expert. All translations were merged into a single version by the committee. The single Malay version was distributed to three independent translators for a back translation process (into English). Comparison between English backward translation and original uMARS was done by the team members and the first Malay Language version of uMARS was produced. The first version was then used for consequent validation. The process of forward-backward translation is demonstrated in Figure 2.
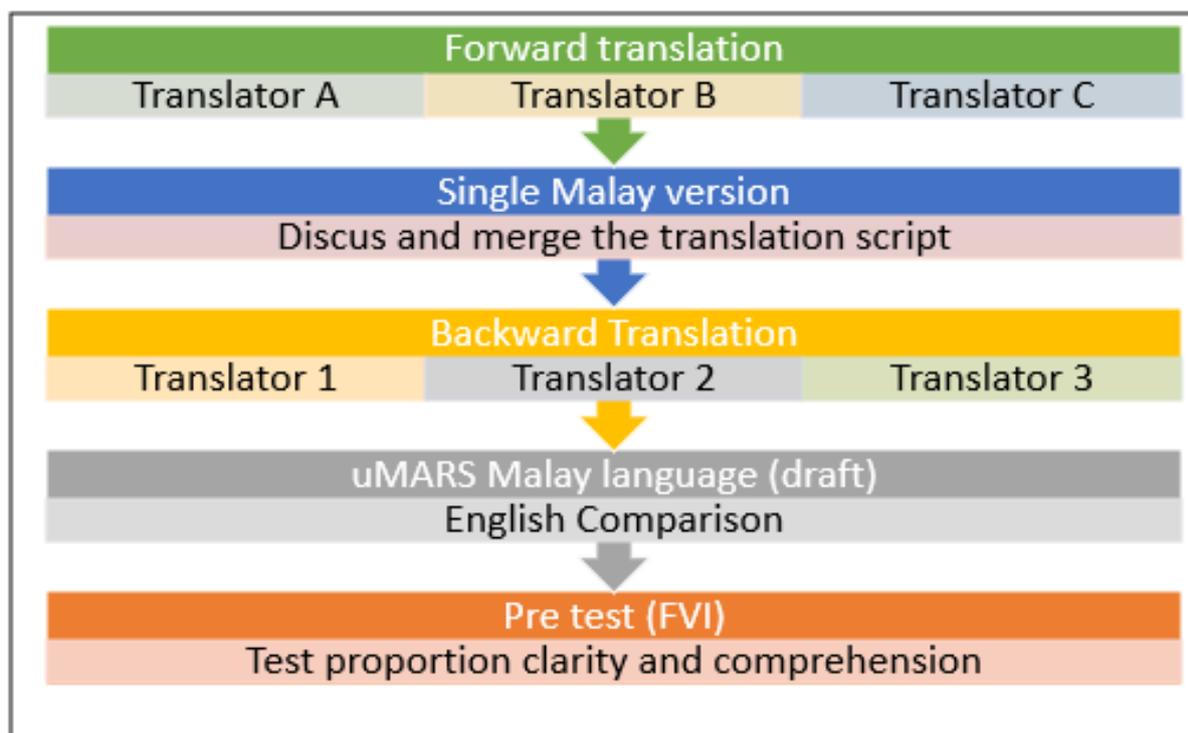

Figure 2. Process forward backward translation.

Assessment of reliability and validity

The face validation process involved ten respondents. Each respondent would remain anonymous to one another during the face validation review procedure, which was carried out independently. The face validaion process was conducted to test proportion clarity and comprehension using the Face Validation Index (FVI)(Yusoff, 2019). uMARS Malay version was distributed to ten respondents among the public, patients, students and lecturers. Data management and data analysis were conducted using Microsoft Excel 2016. The face validity assessment was done using the Face Validity Index (FVI) to measure the degree of clarity and comprehension. FVI is a validity method which is widely used by

researchers because its calculation is straightforward, easily understood and informative (Chin *et al.*, 2018). Respondents who might be potential users were asked to rate each item on a scale of 1 to 4 based on the degree of clarity and comprehension of the item being assessed (1 = not clear and 4 = very clear). Scores 3 and 4 were re-categorized into 1 (clear and understandable) and others into 0 (not clear and understandable). FVI was calculated based on two components which were Item-FVI (I-FVI) and Scale S-FVI (S-FVI). Since this study used ten respondents as raters, the acceptable cut-off score of FVI is at least 0.83 as mentioned by (Marzuki *et al.*, 2018).

Correspondingly, S-FVI is derived from the cumulative scores of items rated as relevant in the tool. The S-FVI can be calculated based on the Universal Agreement (UA) among raters (S-FVI/UA) or based on the average FVI (S-FVI/Ave), with the former being the more conservative method. Calculating the S-FVI/UA is done by adding all items with I-FVI equal to one and dividing it by the total number of items, while S-FVI/Ave can be calculated based on item (I-FVI) and based on proportion clarity and comprehension score across the raters. S-FVI/UA and S-FVI/Ave > 0.8 are considered to represent a satisfactory level of response process validity.

In the validation study for the uMARS Malay version, the sample size was selected to meet statistical power requirements and ensure reliable results. A sample of 36 participants was chosen based on recommendations for achieving adequate internal consistency and validity in cross-cultural adaptation studies. This number aligns with guidelines suggesting a minimum of 30 participants for pilot testing reliability using measures like Cronbach's alpha, which ensures that the tool can be evaluated for consistency across responses (Bujang *et al.*, 2018). Additionally, this sample size provides sufficient power to detect any major issues with item clarity, usability, and overall tool functionality, making it a practical choice for a preliminary validation phase.

The reliability of a questionnaire refers to the degree of consistency achieved while producing the result. The reliability of the questionnaire in terms of its internal consistency was determined using Cronbach's alpha. Cronbach's alpha is an essential statistical tool in research because it ensures that the items in a test or survey are consistently measuring the intended construct, contributing to the overall reliability and validity of the study's findings. The uMARS Malay Version (draft) was piloted with 36 respondents using a mobile health app developed for a community health promotion program (Bujang *et al.*, 2018). The respondents were selected randomly among undergraduate and postgraduate students at the university and individuals who came to the dental clinic either as patients or accompanying others. An information sheet containing pertinent information about the study was given to eligible participants, and informed consent was obtained. The questionnaire also included a section on participant demographic information. Data management and analyses were conducted using SPSS version 26 (IBM). Analysis was done to assess the reliability of items and construct. The summary of Cronbach's alpha score and the internal consistency is illustrated in Table 1.

Result

Demographics

Out of 36 respondents, 8 (22.2%) were males and 28 (77.8%) were females, and they are in the range of 19 years old – 51 years old. 100% are Malay respondents. Regarding the type of users, 11 (30.6%) were healthcare providers, 24 were derived from a pool of patients (66.7%), and 1 was a (2.8%) student. They were asked about their self-rated level of computer skills. The result showed that 20(55.6%) proclaimed to possess a moderate level of computer skills. The majority of the respondents 25 (69.4%), were android users.

Cross cultural adaptation and translation process

No significant discrepancies were derived when comparing the initial version of uMARS to its successor.   Minor changes were carried out to adhere to comments imposed by lay experts; all of which centred on grammar, punctuation, terminology as well as sentence structure. A total of 26 items with five constructs were collectively included to be processed for validation and reliability testing.

Reliability and validity of uMARS Malay Version.

A summary of the entire adaptation and face validation process and findings is illustrated in Table 2.

Validation of Malay Version of uMARS

A summary of the internal reliability findings is illustrated in Table 3.

Cronbach alpha for all constructs is in the range of 0.857 – 0.984 as the result for objective quality domain is 0.918. This result indicated that uMARS Malay version has excellent internal reliability. Eliminating any item would not increase the Cronbach alpha. High corrected item-total correlations indicated significant homogeneity and ranged from 0.65 - 0.98.

Table 1. Cronbach Alpha and internal consistency in dental research proposed by (Jain & Angural, 2017).

Cronbach's alpha Internal consistency	
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

Table 2. The FVI indices.

FVI Indices	Formula	Result
S-FVI/Ave (Based on I-FVI)	(Sum of I-FVI) / Number of items	0.99
S-FVI/Ave (Based on proportion clarity and comprehension)	(Sum of Proportion clarity &comprehension) / Number of items	0.99
S-FVI/UA	(Sum of UA Scores) / Number of items	0.88

Table 3. Internal reliability for Objective Quality uMARS Malay version.

Questions	Mean ± SD	ITC	Cronbach's alpha
A: Engagement			
A1: Entertainment	3.61 ± 1.08	0.83	0.914
A2: Interest	3.78 ± 1.10	0.83	
A3: Customisation	3.33 ± 1.39	0.67	
A4: Interactivity	3.47 ± 1.08	0.86	
A5: Target Group	3.64 ± 1.05	0.77	
B: Functionality			
B6: Performance	3.56 ± 1.10	0.91	0.957
B7: Ease of use	3.71 ± 1.06	0.91	
B8: Navigation	3.78 ± 0.99	0.92	
B9: Gestural Design	3.83 ± 0.97	0.86	
C: Aesthetics			
C10: Layout	3.67 ± 0.93	0.90	0.958
C11: Graphics	3.61 ± 0.93	0.91	
C12: Visual Appeal	3.69 ± 0.95	0.92	
D: Information			
D13: Quality of Information	3.73 ± 0.72	0.72	0.909
D14: Quantity of Information	4.00 ± 0.71	0.81	
D15: Visual Information	4.00 ± 0.71	0.84	
D16: Credibility of source	4.24 ± 0.71	0.81	
Total for Objective Quality			**0.918**
E: App subjective quality			
E17: Recommendation	3.67 ± 1.01	0.67	
E18: Usage	3.03 ± 1.13	0.80	
E19: Interest to pay	2.58 ± 1.16	0.65	
E20: Star rating	3.64 ± 0.93	0.71	
Total for Objective Quality			**0.857**
F: Perceive Impact			
F1: Awareness	3.47 ± 1.34	0.97	
F2: Knowledge	3.47 ± 1.27	0.97	
F3: Attitudes	3.39 ± 1.32	0.98	
F4: Intention to change	3.44 ± 1.32	0.98	
F5: Help seeking	3.53 ± 1.32	0.96	
F6: Behaviour Change	3.56 ± 1.21	0.83	
Total for Perceive Impact			**0.984**

Discussion

The Malay language version of uMARS has thus far proven to be both accurate as well as reliable, the questionnaires of which were further translated into Malay via a standard procedure for cross-cultural adaption to preserve its quality and maintain synchronicity to its predecessor. (Beaton *et al*., 2000). Adapting a questionnaire to meet the requirements of the local context and to fit the local needs are vital steps in ensuring the validity and reliability of the data collected (Sousa *et al*., 2017). Researchers will be able to carry out future studies and assess the quality of mobile health application development by using uMars Malay Versions that are adapted to the local context. It will specifically serve as a guide for developers to strengthen the process of developing apps for planning, managing, and improving the apps and ensuring that the built-in applications can be utilised continually by users to improve their health results. The outcomes showed that the uMARS Malay Version, which was culturally adjusted, is acceptable. Employing rigorous and context-specific translation methods,

along with a combined translation approach, is essential for achieving accurate and culturally relevant adaptations of research tools. This approach helps retain the original meaning while making the tool suitable for the target audience's culture and language, supporting the tool's reliability and validity (Sousa & Rojjanasrirat, 2011). The uMARS Malay version shows a Cronbach's alpha of 0.918 for objective quality and 0.857 for subjective quality, indicating that these scores align well with those found in Arabic language versions that reported internal consistency around 0.88 (Bardus *et al.*, 2020), Cronbach's alpha of 0.75–0.85 for Japanese version (Shinohara *et al.*, 2022), whereas Cronbach's alpha for original version = 0.90 (Stoyanov *et al.*, 2016), showing uMARS Malay Version are comparable reliability across cultural adaptations

This study also presents the results of adaptation, face validation, and internal reliability of the uMARS Malay version. Validity index (FVI) calculations were produced for 26 items, which are five in section A (engagement), four in section B (functionality), three in section C (aesthetics), four in section D (information), four in section E (subjective quality), and 6 in Section F (Perceived impact). The S-FVI/Ave for uMARs Malay version is 0.99, and S-FVI/UA is 0.88, which shows that the uMARS Malay version achieved a satisfactory level (Yusoff, 2019). The Malay uMARS face validity index (S-FVI/Ave) was 0.99, note that it is consistent with scores in Spanish and Japanese versions, which also reported high validity. This consistency shows the thoroughness of the uMARS adaptation (Martin-Payo *et al.*, 2021). It demonstrates that the Malay uMARS meets international standards, supporting its reliability and validity for Malay-speaking users.

The newly produced Malay language version of the application was screened by ten lay experts derived from the public population, from groups of students as well as lecturers that had no prior participation in any study phases. It served as a means of establishing the app's clarity and understandability. The

expert can ensure that the community for which the questionnaire is being designed is accurately represented by employing possible members of the target group as the subject (Zamanzadeh *et al.*, 2015).

Internal consistency was chosen to evaluate the questionnaire based on its widespread use (Bolarinwa, 2015; Taber, 2018) and feasibility. Cronbach's alpha was used to show that the different questions in the questionnaire reliably measured the app quality subscale in the target user. Having the 0.857 – 0.984 for the construct and objective quality domain carried 0.918 as a Cronbach alpha presented with excellent internal reliability for the full scale and good level for the subscale (Jain & Angural, 2017). The Cronbach alpha in this study would definitely meet the criteria if it exceeded 0.7 (Yusoff *et al.*, 2021). All items in the uMARS Malay version in the reliability test executed in the range of 0.65 to 0.98 item-total correlation (ITC), which indicates that all items are highly discriminating. The link between an item and the overall score on the other items is defined by the corrected item-total correlation (Zijlmans *et al.*, 2019). The scale range of uMARS Malay version is proof of the application's accessibility, validity as well as reliability as a means of defining the quality of mobile apps and the perceived impact imposed on mobile apps though some mobile applications were developed using scientific research, it's still critical to evaluate their contributions to health education (Karlsen *et al.*, 2022). It is crucial to bear in mind that user adherence is also influenced by the quality of mobile applications while using them to improve population healthcare.

This study produced a validated questionnaire for the assessment of quality mobile apps by adapting a validated version of uMARS. The development of the uMARS Malay Version has gone through a rigorous translation and validation process involving panels from a wide range of discipline and potential users to help bridge the cultural gap between differing languages. However, this study has some limitations. Firstly, Malay uMARS was developed for mHealth in a specific group which is primary care dental

set-up. Therefore, some testing on a variety of mobile health systems is required. This study only focuses on face validity index. This is because it is regarded as a component of content validity, as face validity is frequently not recognized as a separate validation method. If a questionnaire was administered on its own, the lack of accuracy is necessary to demonstrate its validity. However, combining face validity with other validation tests can be useful (Polit & Beck, 2006). Furthermore, certain nuances may affect the generalizability of findings beyond culturally similar groups in Malaysia, Brunei, Singapore, and parts of Indonesia. The adapted uMARS may be less applicable to Malay speakers in other settings or to other cultural groups, as interpretations of terms or health app usage could vary. To enhance generalizability, future studies could include testing with a more diverse range of Malay-speaking populations or implement additional cultural adaptations.

Conclusion

Employment of this mobile application is sufficient in fortifying the quality of communication in terms of service delivery whilst also ensuring the availability of continuous care. Information material must contain scientifically reliable information and be provided in a form that is acceptable and useful to the patient. It is necessary to educate patients on how to gain quality mobile health apps to adhere users for sustained use and, reach appropriate agreements and adapt to patient health information. Therefore, uMARS Malay version is a valid and reliable tool to help them to measure the quality of the apps.

Future research should expand validation efforts for the Malay uMARS to include a broader demographic, incorporating users from various regions and with different levels of digital literacy. Further studies might also explore the tool's adaptability across other Malay-speaking communities to strengthen its generalizability and address any remaining cultural nuances.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Abbreviation

uMARS : User mobile app rating scale
FVI : Face Validation Index
I-FVI : Item- Face Validation Index
S-FVI : Scale- Face Validation Index
UA : Universal Agreement
S-FVI/UA: Universal Agreement among raters
S-FVI/Ave: Face Validation Index based on the average.

Ethical Approval

Approval was granted by the Research Ethics Committee of Universiti Teknologi MARA (Ref No: REC/06/2021 (MR/400)), (Date 11th June 2021)

Authors' Contributions

Methodology, Tuan Yuswana Tuan Soh, Supervision, Budi Aslinie Md Sabri; Writing – original draft, Tuan Yuswana Tuan Soh; Writing – review & editing, Nik Mohd Nik Mohd Rosy and Budi Aslinie Md Sabri.

References

Barbosa, S. d. F. F., & Marin, H. d. F. (2009). Web-based simulation: a tool for teaching critical care nursing. *Revista Latino-Americana de Enfermagem*, 17, 7-13.

Bardus, M., Awada, N., Ghandour, L. A., Fares, E.-J., Gherbal, T., Al-Zanati, T., & Stoyanov, S. R. (2020). The Arabic version of the Mobile App Rating Scale: development and validation study. *JMIR mHealth and uHealth*, 8(3), e16956.

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.

Bendotti, H., Lawler, S., Ireland, D., Gartner, C., Hides, L., & Marshall, H. M. (2022). What do people want in a smoking cessation app? An analysis of user reviews and app quality. *Nicotine and Tobacco Research*, 24(2), 169-177.

Berhanuddin, M. A., Mohamad, N. F., & Hassan, E. H. (2017). Development and evaluation of an evidence based smoking cessation app for the Malaysian population: the self-determination theory approach. *ESTEEM Academic Journal*, 13, 28-39.

Bujang, M. A., Omar, E. D., & Baharum, N. A. (2018). A review on sample size determination for Cronbach's alpha test: a simple guide for researchers. *The Malaysian Journal of Medical Sciences*, 25(6), 85.

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *Journal of Advanced Nursing*, 58(4), 386-395.

Chin, R. W. A., Chua, Y. Y., Chu, M. N., Mahadi, N. F., Wong, M. S., Yusoff, M. S., & Lee, Y. Y. (2018). Investigating validity evidence of the Malay translation of the Copenhagen Burnout Inventory. *Journal of Taibah University Medical Sciences*, 13(1), 1-9.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1), 60-95.

Delone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*, 19(4), 9-30.

Heifetz, M., & Lunsky, Y. (2018). Implementation and evaluation of health passport communication tools in emergency departments. *Research in Developmental Disabilities*, 72, 23-32. https://www.sciencedirect.com/science/article/abs/pii/S089142221730255X?via%3Dihub

Hernández, A., Hidalgo, M. D., Hambleton, R. K., & Gómez Benito, J. (2020). International test commission guidelines for test adaptation: A criterion checklist. *Psicothema*, 32 (3), 390-398.

Jain, S., & Angural, V. (2017). Use of Cronbach's alpha in dental research. *Medico Research Chronicles*, 4(3), 285-291.

Karlsen, I. L., Svendsen, P. A., & Abildgaard, J. S. (2022). A review of smartphone applications designed to improve occupational health, safety, and well-being at workplaces. *BMC Public Health*, 22(1), 1-13.

Lu, D. J., Girgis, M., David, J. M., Chung, E. M., Atkins, K. M., & Kamrava, M. (2021). Evaluation of mobile health applications to track patient-reported outcomes for oncology patients: a systematic review. *Advances in Radiation Oncology*, 6(1), 100576.

Mang, J. M., Seuchter, S. A., Gulden, C., Schild, S., Kraska, D., Prokosch, H.-U., & Kapsner, L.A. (2022). DQAgui: a graphical user interface for the MIRACUM data quality assessment tool. *BMC Medical Informatics and Decision Making*, 22(1), 1-11.

Martin-Payo, R., Carrasco-Santos, S., Cuesta, M., Stoyan, S., Gonzalez-Mendez, X., & Fernandez-Alvarez, M. d. M. (2021). Spanish adaptation and validation of the User Version of the Mobile Application Rating Scale (uMARS). *Journal of the American Medical Informatics Association*, 28(12), 2681-2686.

Marzuki, M. F. M., Yaacob, N. A., & Yaacob, N. M. (2018). Translation, cross-cultural adaptation, and validation of the Malay version of the system usability scale questionnaire for the assessment of mobile apps. *JMIR Human Factors*, 5(2), e10308.

Morselli, S., Sebastianelli, A., Domnich, A., Bucchi, C., Spatafora, P., Liaci, A., *et al.* (2021). Translation and validation of the Italian version of the user version of the Mobile Application Rating Scale (uMARS). *Journal of Preventive Medicine and Hygiene*, 62(1), E243.

Park, H. S., Cho, H., & Kim, H. S. (2016). Development of a multi-agent m-health application based on various protocols for chronic disease self-management. *Journal of Medical Systems*, 40(1), 36.

Perez-Jover, V., Sala-González, M., Guilabert, M., & Mira, J. J. (2019). Mobile apps for increasing treatment adherence: systematic review. *Journal of Medical Internet Research*, 21(6), e12505.

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, 489-497.

Pretlow, R., Stock, C., Allison, S., & Roeger, L. (2015). Treatment of child/adolescent obesity using the addiction model: a smartphone app pilot study. *Child Obesity*, 11 (3), 248–259.

Price, P. C., Jhangiani, R. S., & Chiang, I.-C. A. (2015). Reliability and validity of measurement. *Research Methods in Psychology*.

Sereda, M., Smith, S., Newton, K., & Stockdale, D. (2019). Mobile apps for management of tinnitus: users' survey, quality assessment, and content analysis. *JMIR mHealth and uHealth*, *7*(1), e10353.

Shao, K., Huang, J., Song, H., Li, R., & Wu, J. (2014). DAYA: a system for monitoring and enhancing children's oral hygiene. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems* (pp. 251-256).

Sharma, H., Suprabha, B. S., & Rao, A. (2021). Teledentistry and its applications in paediatric dentistry: A literature review. *Pediatric Dental Journal*, 31(3), 203-215.

Sharpe, J. D., & Kamara, M. T. (2018). A systematic evaluation of mobile apps to improve the uptake of and adherence to HIV pre-exposure prophylaxis. *Sexual Health*, 15(6), 587-594.

Shinohara, Y., Yamamoto, K., Ito, M., Sakata, M., Koizumi, S., Hashisako, M., *et al.* (2022). Development and validation of the Japanese version of the uMARS (user version of the mobile app rating system). *International Journal of Medical Informatics*, *165*, 104809.

Soler, C., Zacarías, A., & Lucero, A. (2009). Molarcropolis: a mobile persuasive game to raise oral health and dental hygiene awareness. Proceedings of the International Conference on Advances in Computer Enterntainment Technology

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.

Sousa, V.E., Matson, J., & Dunn Lopez, K. (2017). Questionnaire adapting: little changes mean a lot. *Western Journal of Nursing Research*, *39*, 1289-1300. (0193-9459)

Stoyanov, S. R., Hides, L., Kavanagh, D. J., & Wilson, H. (2016). Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR mHealth and uHealth*, 4(2), e5849.

Strodl, E., Shakespeare-Finch, J., Alichniewicz, K. K., Brown, K., Quinn, C., Hides, L., White, A., Gossage, G., Poerio, L., & Batras, D. (2020). Clinicians' perceptions of PTSD coach australia. *Internet Interventions*, 21, 100333.

Subramani Parasuraman, A. T. S., Yee, S. W. K., Chuon, B. L. C., & Ren, L. Y. (2017). Smartphone usage and increased risk of mobile phone addiction: A concurrent study. *International Journal of Pharmaceutical Investigation*, 7(3), 125.

Underwood, B., Birdsall, J., & Kay, E. (2015). The use of a mobile app to motivate evidence-based oral hygiene behaviour. *British Dental Journal*, 219(4), E2-E2.

Yusoff, M. S. B. (2019). ABC of response process validation and face validity index calculation. *Education in Medicine Journal*, 11(10.21315)

Yusoff, M. S. B., Arifin, W. N., & Hadie, S. N. H. (2021). ABC of questionnaire development and validation for survey research. *Education in Medicine Journal*, 13(1).

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.

Zhang, X., Chou, J., & Wang, F. (2018). Integrative analysis of patient health records and neuroimages via memory-based graph convolutional network. 2018 IEEE International Conference on Data Mining (ICDM)

Zijlmans, E. A., Tijmstra, J., Van der Ark, L. A., & Sijtsma, K. (2019). Item-score reliability as a selection tool in test construction. *Frontiers in Psychology*, 9, 2298.