Test-Retest Reliability of Malay Version of Stroke Knowledge Test

Siti Noorkhairina Sowtali[1], Norhasyimah Ghazali[1], Dariah Mohd. Yusoff[2]*, Sakinah Harith[3], Monniaty Mohamed[4], Nur 'Ain Ishak[5]

[1]*Kulliyyah of Nursing, International Islamic University Malaysia, Level 2, Jalan Hospital Campus, Kuantan, Pahang Darul Makmur, Malaysia.*
[2]*School of Health Sciences, Universiti Sains Malaysia, 16150 KubangKerian, Kelantan, Malaysia.*
[3]*Faculty of Health Sciences, Universiti Sultan Zainal Abidin, 21300 Kuala Terengganu, Terengganu, Malaysia.*
[4]*Medical Department, Hospital Raja Perempuan Zainab II, 15000, Kota Bharu, Kelantan, Malaysia.*
[5] *Subang Jaya Medical Centre, Jalan SS 12/1a, SS1, 47500 Subang Jaya, Selangor, Malaysia.*

ABSTRACT

Objective: This study was aimed to evaluate the stability and reliability of the Malay version of the Stroke Knowledge Test among kidney disease patients at the Hemodialysis Unit and Nephrology Clinic of Hospital Tengku Ampuan Afzan, Kuantan. **Methods**: A cross-sectional study was conducted on 30 conveniently-selected kidney disease patients between April and May 2017. The questionnaires consisted of three parts: sociodemographic background, clinical, and lifestyle-related risk factors; the Malay version of the Stroke Knowledge Test were administered on the patients in a pre-test/post-test manner with a gap of two weeks. **Results**: The Malay version of the Stroke Knowledge Test had a fair pre-test/post-test agreement (ICC=0.247) when administered on the kidney disease patients to measure their knowledge of stroke. The items were considered to have acceptable difficulty and discrimination indexes. They were also reliable as all aspects of stroke were covered. **Conclusions**: The Malay version of the Stroke Knowledge Test had a fair pre-test/ post-test agreement when conducted on the kidney disease patients to check its stability and reliability. The results reflected a mixture of easy and moderate content, as well as acceptable modifications of the items. Hence, as per previous researches, this instrument was considered to be a reliable tool for evaluating the knowledge of stroke outcomes in kidney disease patients. Future larger research is needed to confirm the test-retest reliability of the instrument.

KEYWORDS: Stroke Knowledge Test, reliability, difficulty index, discrimination index, test-retest reliability

INTRODUCTION

Several studies on the association of kidney failure with cardiovascular disease (CVD) and uncontrolled hypertension (HTN) have been performed (1-3). Most of the morbidities and mortalities associated with chronic kidney disease (CKD) are significantly attributable to cardiovascular outcomes, which highlight the fact that CKD and CVD are very much interrelated. Some complications of kidney disease - such as HTN, high homocysteine level, and anemia - can lead to CVD (1-3). Furthermore, in the worst-case scenario, most of these complications will end up causing stroke.

Stages 1 and 2 of CKD constitute the earlier stages of the disease, meanwhile stage 3 is the moderate stage. Stages 4 and 5 denote a late stage whereby the kidney functions have significantly deteriorated and so the patient requires renal replacement therapies (RRT) such as hemodialysis (HD), continuous ambulatory peritoneal dialysis (CAPD), peritoneal dialysis (PD), and kidney transplantation (1-4). Sometimes, the term CKD stage 5 is also used interchangeably with kidney failure or end-stage kidney disease/ failure (ESRD/ ESRF) in Malaysia (3,4). The treatment modalities for CKD include the optimization of blood pressure and glycemic control, reduction of proteinuria, as well as control of dyslipidemia to reduce cardiovascular events (3,4).

Approximately 1,155 people per million have undergone dialysis (either HD or PD), while 61 per million have undergone kidney transplantation in 2014 (3,4). With more than 5,000 new patients diagnosed with kidney failure every year, the number of Malaysians who were dependent on dialysis has increased to more than 30,000 in 2015. The Malaysian Ministry of Health (MOH) and other non-government organizations (NGOs) have organized many public-educational activities to create awareness of kidney disease. A kidney disease patient is at high risk for CVD. Changes in dietary habits as well as lack of physical activity/ exercise may also lead to vascular diseases such as stroke in the long run (5).

The current World Health Organization (WHO) definition of stroke has been introduced in 1970 and is still in use to date (6). It is defined as "rapidly-developing clinical signs of focal (or global) disturbance of cerebral function, lasting more than 24 hours or leading to death, with no apparent cause other than that of vascular origin" (6). According to previous finding, Asian populations had a higher incidence of stroke and mortality from the same; the values being 116 and 483/100,000 per year respectively (7). This is an alarming occurrence since information on the incidence and prevalence of stroke in Malaysia is still lacking owing to registry matters (7,8). This highlights the need for creating public awareness

* Corresponding Author

Dariah Mohd. Yusoff
School of Health Sciences,
Universiti Sains Malaysia, 16150,
Kubang Kerian, Kelantan, Malaysia.
Email: dariah@usm.my
Tel: 017-9646477

on the importance of practicing healthy lifestyles in order to manage the medical risk factors that might contribute to the occurrence of stroke. The medical risk factors that are associated with stroke include HTN, diabetes mellitus (DM), dyslipidemia, cardiovascular disease (e.g. atrial fibrillation), and kidney disease. Furthermore, having knowledge on stroke area may prevent someone from both primary and recurrent attack.

Therefore, in order to evaluate the association of the stability and reliability of stroke related instrument on the public or patients, previous researcher have translated the Stroke Knowledge Test (SKT) into Malay and validated the instrument on stroke patients in one of the government hospitals in Kelantan, Malaysia (9). This instrument was originally developed by researchers in Australia, where it was administered on the stroke survivors and caregivers who attended a community-based stroke educational program (10,11). The original tool had a Cronbach's alpha of 0.76. Meanwhile, the reliability of the Malay version of the SKT among the stroke patients at USM Hospital was of a Kuder Richardson (KR20) value of 0.58 – a moderate reliability for a new instrument (9).

However, the reliability of the Malay version of SKT may change when administered on a different population due to difference in socio-demographic background (e.g. education level, income status, demographic location) and sample size number (e.g. smaller or larger). Therefore, the aim of this study was to test and retest the stability as well as reliability of the Malay version of SKT. The study was conducted on a different population of kidney disease patients who were easily contactable in order to obtain a more accurate result on the stability of the instrument. Besides, it was hoped that this Malay version of SKT would be able to contribute to the assessment of public education or specifically, measurement of the level of stroke knowledge in the public.

Table 1: The stages of CKD based on GFR

Stage of CKD	Description	GFR (mL/min per 1.73 m²)
1	Normal or increased GFR, with other evidence of kidney damage	More than 90
2	Slight decrease in GFR, with other evidence of kidney damage	60-89
3A	Moderate decrease in GFR, with or without other evidence of kidney damage	45 - 59
3B		30 - 44
4	Severe decrease in GFR, with or without other evidence of kidney damage	15-29
5	Established kidney failure	Less than 15

(Clinical Practice Guidelines Management of Chronic Kidney Disease in Adults, 2011)

METHODS

This was a cross-sectional study. The source population was 30 kidney disease patients who were followed-up or undergoing regular hemodialysis at the Nephrology Clinic and Hemodialysis Unit of Hospital Tengku Ampuan Afzan (HTAA), Kuantan. The sample size was calculated using the Intraclass Correlation Coefficient (ICC) Sample Size Calculator, which was adapted from the Microsoft Excel file prepared by Wan Nor Arifin (2017) (12). The significance level (α) was set at 0.05, power (1-β) at 0.80, acceptable reliability (P0) of 0.58, expected reliability (P1) of 0.80 with drop-out of 20.0% and the required sample size was 23. The calculated value was rounded up to 30 so as to achieve the minimum sample size according to the central limit theorem of normal distribution. This study has adopted a smaller sample size since the respondents were only providing preliminary findings. Thus, precautions were taken against generalizing these findings to the overall CKD patient population in Malaysia since convenience sampling was applied in a single hospital setting.

Ethical clearance was obtained from the Ethical Committees of the International Islamic University of Malaysia (IREC) and Ministry of Health (MREC). Meanwhile, approval for the usage of the study site was obtained from the Director and Head of Nephrology Department via the Clinical Research Centre (CRC) of Hospital Tengku Ampuan Afzan [Ref. no. NMRR-17-131-34139]. The patients were included in this study if they were (1) diagnosed with kidney disease, (2) aged 18 years and above, (3) able to understand Malay language, as well as (4) willing to participate in the research. The data and information were collected over two (2) months. Prior to data collection, the purpose of the study was explained to the respondents. Their confidentiality was assured, and consent was obtained from them.

The questionnaire was self-administered; the patients were given approximately 20 minutes to answer all three (3) parts. Part A was on the sociodemographic data of the participants, Part B medical status and lifestyle habits of the participants, as well as Part C on their level of knowledge regarding the development of stroke. This study utilized the validated Malay version of the SKT from the earlier study on stroke patients (9). The questionnaire consisted of 20 items on the risk factors, signs and symptoms, prevention, prevalence, treatment, as well as rehabilitation of stroke (9). The 20 multiple choice questions (MCQs) contained five options each. Of the five, there was one correct option, three distracters, and an "I don't know" option to avoid guessing. One (1) mark was given for each correct answer and no (0) mark for a wrong answer. The range of possible scores was zero to 20, whereby a higher score indicated better knowledge.

The same participants were approached after two (2) weeks and the same procedure repeated to minimize the learning effect (13). A duration of two weeks was chosen because it was recommended by past studies that the difference in time should be large enough so that the patients were not likely to remember or be influenced by the first set of responses when the second set of questionnaires was given (14,15). Data from Part A and Part B of the questionnaire was analyzed using descriptive statistics. Meanwhile, for Part C, an item analysis was performed along with the determination of the Intraclass Correlation Coefficient (ICC).

Item Analysis: Difficulty Index (P) and Discrimination Index (D)

Item analysis is a valuable and simple post-test procedure as it provides information on the reliability and validity of a test item (16). Furthermore, the item analysis enables the identification of good multiple choice questions (MCQs) based on the difficulty index, discrimination index, and distractor efficiency (16). A good item should be able to assess the

cognitive, affective, as well as psychomotor domains. Such items are preferred over other approaches owing to their (1) objectivity in assessments, (2) comparability in a variety of settings, (3) wide coverage of a subject, and (4) minimal assessor bias (16).

The difficulty index described the percentage of patients who correctly answered an item, and this indicated the difficulty of the item. Meanwhile, the discrimination index distinguished between two groups of patients ('High Mark' and 'Low Mark' groups) who have correctly answered a question (13). After the researcher has summed up the scores for the 20 questions, the sets of answers in Part C were then arranged in descending order. Then, 27.0% of all respondents were extracted from the 'High Mark' group and 'Low Mark' group respectively. A value of 27.0% was used because it could maximize the differences in a normal distribution while providing ample cases for data analysis (17,18). The calculations for the difficulty index and discrimination index of the 20 items were adapted from an earlier study on stroke patients by Siti Noorkhairina et al. (2016) (9).

Intraclass Correlation Coefficient (ICC)

The test-retest reliability (also known as stability or reproducibility) concerned the consistency of scores across two (2) separate measurements over time. Intraclass correlation coefficient (ICC) was used as a tool for evaluating the reliability and stability of the Malay version of SKT. The recommended reliability measure for test-retest reliability is the ICC (19,20). Test-retest reliability has sometimes been reported as a correlation between the pre-test and post-test scores. Furthermore, the ICC was a more desirable measure of reliability since it reflected both the degree of correlation and agreement between the measurements (20).

There were 10 forms of ICCs, which were based on the "Model" (1-way random effect, 2-way random effects, or 2-way fixed effects), "Type" (single rater, measurement, or the mean of k raters/ measurements), and "Definition" of the relationship (consistency or absolute agreement) (20). The ICC used in this study was the 1-way random effect model with a single measure [ICC (1)], which focused on the correlation between two measurements taken at different times (regardless of the number of raters) (21,22). The interpretations of the ICC and kappa coefficient were in accordance with the kappa coefficient guidelines: (a) <0.20: poor agreement; (b) 0.21–0.40: fair agreement; (c) 0.41–0.60: moderate agreement; (d) 0.61–0.80: good agreement, and (e) 0.81–1.00: very good agreement (23).

RESULTS

Sociodemographic Data of kidney disease patients

Overall, 30 kidney disease patients have been recruited to check the stability of the SKT in this study. Table 2 describes the sociodemographic profiles of the kidney disease patients who were registered under the Nephrology Department of Hospital Tengku Ampuan Afzan (HTAA), Kuantan. More than half of the respondents were males (66.7%), with the mean age being 41.9 ± 16.02 years. The majority of them were Malays (96.7%), secondary school finishers (56.7%), married (50.0%), and self-employed (33.3%).

Table 2. Socio-demographic characteristics of the participants (n=30)

Variables	Frequency (n)	Percentage (%)
Gender		
Male	**20**	**66.7**
Female	10	33.3
Age (categorical)		
18 - 30	**12**	**40.0**
31 - 40	1	3.3
41 - 50	4	13.3
51 - 60	10	33.3
> 61	3	10.0
Race		
Malay	**29**	**96.7**
Chinese	1	3.3
Marital Status		
Single	13	43.3
Married	**15**	**50.0**
Widower / widow	2	6.7
Working Status		
Government employee	4	13.3
Private sector worker	3	10.0
Retired / Pension	5	16.7
Self-employed	**10**	**33.3**
Not working	8	26.7
Educational Status		
Primary school	5	16.7
Secondary school	**17**	**56.7**
Certificate / Diploma	5	16.7
Degree / Master / Phd	3	10.0

Note: *SD: Standard Deviation

Clinical and Lifestyle-related Risk Factors

Table 3 highlights the clinical and lifestyle-related risk factors of the kidney disease patients in this study. In terms of the former, the majority of the patients were diagnosed with CKD stage 5 (70.0%), with some undergoing dialysis owing to the development of end stage renal disease (ESRD). More than half had HTN (60.0%), while one in six had DM (16.7%). As for lifestyle risk factors, only 6.7% were cigarette smokers while 93.3% were non -smokers. With reference to the physical activity status of the participants, 33.3% claimed to be involved in such activities at their workplaces.

Table 3. Clinical and lifestyle risk factors (n=30)

Variables	Frequency (n)	Percentage (%)
Stages of kidney disease		
CKD - 3	1	3.3
CKD - 4	5	16.7
CKD - 5 (Undergoing PD)*	3	10.0
CKD - 5 (Undergoing HD)#	**21**	**70.0**
Medical history		
Hypertension	**18**	**60.0**
Diabetes	5	16.7
Anemia	2	6.7
Bone disease	3	10.0
Heart disease	2	6.7
Type of treatment		
Hemodialysis	**21**	**70.0**
Peritoneal dialysis	9	30.0
Smoking status		
Yes	2	6.7
No	**28**	**93.3**
Physical activity status		
Activities in workplace	**10**	**33.3**
Recreation / sports	5	16.7
Walking (daily movement)	3	10.0
Performing house chores	8	26.7
Sitting / lying	4	13.3

Note:*PD = Peritoneal Dialysis; #HD = Hemodialysis

Item Analysis of the Malay version of Stroke Knowledge Test (SKT)

Table 4 shows the results of the item analysis, which included the difficulty index and discrimination index of all 20 items of the Malay version of the Stroke Knowledge Test (SKT) created by Siti Noorkhairina et al. (2016) (9). The post-test (T2) difficulty and discrimination indexes were calculated as these reflected the outcomes of the test. Of the 20 items, the majority had a difficulty index of 0.2 – 0.8, or were easy-to-moderate, except for item 14. Similarly, a few items had low-to-poor discrimination indexes, such as items 5, 6, 7, 8, 9, 11, and 20. It was recommended items with poor or no discrimination and high difficulty be revised in future studies to enable better understanding by the CKD patients.

Stability and Reliability of Stroke Knowledge Test (SKT)

The Intraclass Correlation Coefficient (ICC) in the form of a random one-way single-measure model, ICC (1), was used to check the stability of the Malay version of SKT. The value if the ICC (1) was 0.247, which indicated a fair agreement between the pre-test and post-test results for a single-performance rating and a researcher. The reliability of the item was strong (r = 0.611). As the 95% limits of agreement (LOA) in this case included zero (-6.864, 3.064), these two measures were likely to have a fair agreement over time and be useable in further studies (Figure 1).

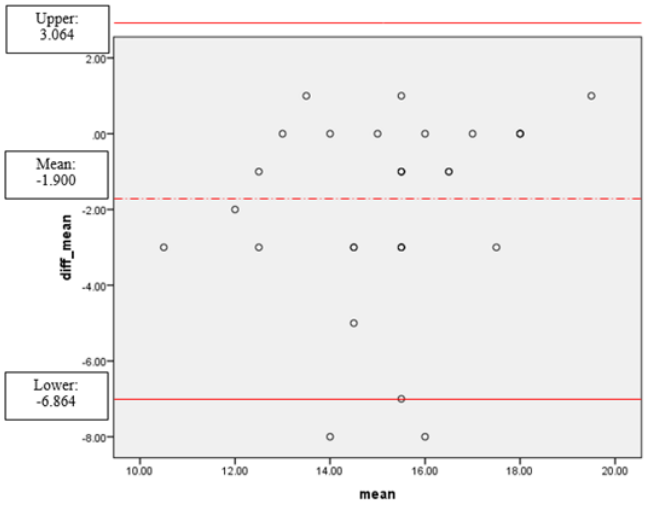

Figure 1. Bland Altman Plot (difference mean two measurement) – pre and post test scores Malay version SKT

DISCUSSION

In this preliminary study, the majority of kidney disease patients were adult males. At this age, several common metabolic disorders might set in and complicate one's kidney functions (1-3). This is particularly true when co-morbidities (such as HTN and DM) and lifestyle factors (like smoking) are present, which strongly predisposes the individual to kidney problems. Although most participants have claimed that they performed physical activities at their workplaces, the authors have not asked for details on the types and durations of the same. There was still ambiguity in the interpretation of the results, so a specific study on the lifestyles of kidney disease patients would be required.

Commonly, married people are more concerned about their health and are more likely to seek treatment when they are not well. This is reflected by the finding that half of the patients were married and had multiple roles in their lives. Among the reasons for the increased health-consciousness of married people is that their role as a spouse or parent requires them to become a breadwinner of their families. People who are working, especially males, need to maintain their health and well-being for an optimal job performance. On another matter, more than half of the kidney disease patients in the study have completed their secondary education. This highlighted the contribution of education to the creation of awareness on the association between kidney problems and stroke, as proven by the item analysis, reliability, and stability of the results in this study. A research in the United Kingdom (UK) has highlighted that sociodemographic factors such as income status and lifestyle patterns did contribute to the occurrence of CKD (24).

Reliability refers to the consistency and repeatability of an instrument over time (27, 28). Although reliability-testing is necessary, it is not a self-sufficient component of instrument validity. This is because the reliability may change when a different population is employed. In the earlier Malay version of the SKT, the researchers have used KR-20 to establish the internal consistency of the test in stroke patients at Hospital USM (24, 25). Thus, it was of the interest of the researchers to evaluate the test-retest reliability of the Malay version of SKT – which was developed by Siti Noorkhairina et al. (2016) – for the presence of changes in the cognitive and trait scales over time (9, 26).

The minimum duration for the evaluation of test-retest reliability is two weeks to six months. However, the latter is not appropriate for states that are expected to change over time, such as attitudes, mood, or knowledge levels following an intervention (26). Therefore, this study has assessed the test-retest reliability over two weeks since the kidney disease patients were on regular follow-up sessions at HTAA. The results of the reliability test were rated as moderate when conducted on the kidney disease patients; this was acceptable according to Downing and Haladyna (2011) as well as Sullivan and Dunton (2004) (27, 10). The results suggested that the Malay version of SKT could be repeated on other patients with different comorbidities of stroke such as hypertension, diabetes, dyslipidemia, and metabolic disorder – all of which are linked with the occurrence of cardiovascular diseases as well.

Item analysis of SKT showed that most items (>50%) had acceptable values for the difficulty and discrimination indexes. In terms of difficulty index, 70% of the items were found to be easy, 25% moderate, and only 5% difficult. As for the discrimination index, 55% of the items were excellent-to-acceptable while 45% poor-to-no discrimination. In comparison with the initial findings of Siti Noorkhairina et al. (2016), the ranges of both difficulty and discrimination indexes in this study were wider, which reflected a better understanding of stroke (9). It could be that kidney disease patients with multiple risk factors had better knowledge of stroke.

Interestingly, only one item (5%) of 20 in this study was rated to be difficult by the majority of kidney disease patients. This item was on the epidemiology of stroke in Malaysia. A similar outcome was found in Siti Noorkhairina et al. (2016) as well as Sullivan and Dunton (2004) (9-10). This is a common

occurrence because epidemiology is not of interest to the general public, even though it is an important form of knowledge for healthcare professionals. Overall, the findings of the item analysis highlighted that there was good awareness that stroke was one of the risk factors of kidney disease. Nevertheless, a larger study should be conducted to confirm this.

In terms of the stability and reliability of the Malay version of the Stroke Knowledge Test (SKT), the Intraclass Correlation Coefficient (ICC) denoted a fair pre-test/ post-test agreement (ICC = 0.247) when the instrument was administered on the kidney disease population. In the concept of ICC, there are no standard values for acceptable reliability (17). However, a low ICC might not only reflect a poor rater or measurement agreement, but also a (1) lack of variability among the subjects, (2) small number of subjects, and (3) small number of raters (26). An alternative method to determine the absolute measure of agreement between two (2) measurements of the same instrument is by visualization through the Bland-Altman plot, which is commonly used to evaluate reliability (17).

Based on the Bland-Altman plot, it could be justified that the Malay version of SKT had only a fair pre-test/ post-test agreement when conducted on different populations, unlike that previous finding (9). Nevertheless, this study had its own limitation. Researcher would like to highlight that 'Hawthorne' effect (learning effect) may occur during this study since it involved repeated measurement. For an instance, the participant may perform better at the post-test due to memorizing the questions or having first experience answering the questions and become familiar with the answer.

CONCLUSION

Although the reliability and stability of the Malay version of SKT had a fair pre-test/ post-test agreement and moderate ICC, the quality of the results can be improved by tests on larger populations of CKD patients. It is believed that doing so will help determine the presence of significant effects on the item analyses of the SKT. Besides, further test-retest reliability assessments on different but directly-reachable populations are needed to prove and support the reliability as well as stability which have been reported by previous researchers.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest in this study.

ACKNOWLEDGEMENT

The authors would like to thank the International Islamic University Malaysia (IIUM) for providing funding via the Research Incentive Grant Scheme (RIGS 16-143-0307). Our gratitude also goes to the team members of the RIGS grant (from Universiti Sains Malaysia and Universiti Sultan Zainal Abidin), multidisciplinary healthcare professionals of IIUM Medical Center (IIUM MC) and Hospital Tengku Ampuan Afzan (HTAA), as well as the research staff at the Clinical Research Center (HTAA). Special thanks to Miss Lim Bee Chiu for the statistical analysis consultations, and to those who have directly or indirectly assisted us in this endeavor.

REFERENCES

1. Malaysia Ministry of Health. (2017). Clinical Practice Guideline: Management of Hypertension, 4th Edition. Retrieved [Accessed 19012018].
2. Malaysia Ministry of Health. (2017). Primary and Secondary Prevention of Cardiovascular Disease [Accessed 19012018].
3. Malaysia Ministry of Health. (2011). Clinical Practice Guideline: Management of Chronic Kidney Disease in Adults. Retrieved [Accessed 19012018].
4. Malaysia Ministry of Health. (2014). 22nd Report of the Malaysian Dialysis and Transplant Registry. Retrieved [Accessed 01022018].
5. Siti Noorkhairina S, Sakinah H, CheRabiaah M. Secondary Stroke Prevention through Patient Education Intervention on Lifestyle Risk Factors: A Review. Health and the Environment Journal,2013; Vol 4(2):127-51.
6. Sacco RL, Kasner SE, Broderick, JP, et al. An Updated Definition of Stroke for the 21st Century: A Statement for Healthcare Professionals from the American Heart Association/American Stroke Association. American Heart Association. Stroke. 2013; 44:2064-89.
7. Suwanwela NC, Poungvarin N. Stroke burden and stroke care system in Asia. Neurology India. 2016; 64:46□51.
8. Tharakan, J. (2012). Stroke Registry-Relevance and Contributions. *Med J Malaysia,* **67(3)**,251.
9. Siti Noorkhairina S, Yusoff DM, Harith S. Translation and Validation of the Malay version of the Stroke Knowledge Test. Journal of Arrhythmia 2016; 32: 112-118.
10. Sullivan K, Dunton NJ. Development and validation of the Stroke Knowledge Test. Top Stroke Rehabil 2004;11: 19–28.
11. Sullivan KA, Waugh D. Stroke knowledge and misconceptions among survivors of stroke and a non-stroke survivor sample. Top Stroke Rehabil. 2005;12(2):72-81.
12. Arifin WN. Sample size calculator (Version 2.0) [Spreadsheet file]. Author: 2017. Available from http://wnarifin.github.io
13. Hingorjo MR, Jaleel F. Analysis of One-Best MCQs: the Difficulty Index, Discrimination Index and Distractor Efficiency. J Pak Med Assoc 2012; 62 (11): 142-147.
14. Tobler-Ammann BC, de Bruin ED, Fluet M-C et al. Concurrent validity and test-retest reliability of the Virtual Peg Insertion Test to quantify upper limb function in patients with chronic stroke. Journal of NeuroEngineering and Rehabilitation 2016; 24(8):1-14.
15. Marx RG, Menezes A, Horovitz L et. al. A comparison of two time intervals for test-retest reliability of health status instruments. Journal of Clinical Epidemiology. 2003; 56:730-5.
16. Gajjar S, Sharma R, Kumar P et al. Item and Test Analysis to Identify Quality Multiple Choice Question (MCQs) from an Assessment of Medical Students of Ahmedabad, Gujarat. Indian J of Community Med. 2014; 39(1):17-20.
17. Wiersma W, Jurs, SG). Educational measurement and testing. 2nd ed. Boston, MA: Allyn and Bacon; 1990. Retrieved from www.clr.ui.ac.id/wp-content/uploads/.../Educational-Measurement-and-Testing.pdf.
18. Ebel RL, Frisbie DA. Essentials of educational measurement. 5th ed. NJ:

Prentice Hall; 1991.
19. DeVet HCW, Terwee C, Mokkink LB et al. Measurement in Medicine: A practical guide. *Cambridge: Cambridge University Press*; 2011.
20. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Med.2015; 15(2):155-163.
21. McGraw KO, Wong S. Forming inferences about intraclass correlation coefficients. Psychological methods. 1996; 1(1):30-46.
22. Vaz S, Falkmer T, Passmore AE, et al. The Case for Using the Repeatability Coefficient When Calculating Test-Retest Reliability. PLoS ONE. 2013; 8(9): e73990.
23. McHugh ML. Interrater reliability: the kappa statistic. 2012; 22(3): 276-82.
24. Bello AK, Peters J, Rigby J. Socioeconomic Status and Chronic Kidney Disease at Presentation to a Kidney Service in the United Kingdom. Clin J Am SocNephrol. (2008). 3(5): 1316-23.
25. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. Am J Health Syst Pharm 2008;65:2276–84.
26. DeVon HA, Block ME, Moyle-Wright P, et al. A psychometric toolbox for testing validity and reliability. J NursScholarsh 2007;39:155–64.
27. Downing SM, Haladyna TM. Handbook of test development. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers; 2011.
28. Portney LG, Watkins MP Foundations of Clinical Research: application to practice, 3rd ed. NJ: Pearson Education Inc. 2009.

No.	Item in Malay (English)		Difficulty Index	Discrimination Index
1	*Strok yang kerap berlaku di dalam masyarakat Malaysia berpunca daripada* (The most common type of stroke occurs when)		0.88	0.25
		Interpretation	Easy	Acceptable
2	*Faktor berikut meningkatkan risiko anda diserang strok sebanyak dua kali ganda?* (Which of the following will double your risk of stroke?)		0.75	0.5
		Interpretation	Easy	Excellent
3	*Sejenis keadaan degupan jantung laju dan tidak teratur dipanggil Fibrilasi Atria (AF) akan* (A type of irregular heartbeat known as atrial fibrillation (AF))		0.63	0.5
		Interpretation	Moderate	Excellent
4	*Manakah antara peringkat umur berikut **PALING BERISIKO** mengalami strok?* (Which age group is more at risk of stroke?)		0.56	0.875
		Interpretation	Moderate	Excellent
5	*Tanda-tanda amaran kejadian serangan strok kecil (TIA) akan hilang* (The warning signs of transient ischemic attack (TIA) disappear)		0.94	0.125
		Interpretation	Easy	Poor
6	*Antara berikut, manakah tanda-tanda awal serangan strok?* (Which of the following is a warning sign of stroke?)		0.44	0.125
		Interpretation	Moderate	Poor
7	**Tujuan rehabilitasi (pemulihan) bagi individu yang mengalami strok ialah** **(For someone who has had a stroke, the main purpose of rehabilitation is to)**		1.00	0
		Interpretation	Easy	No discrimination
8	**Pengambilan ubatan Aspirin boleh mengurangkan serangan strok secara** **(Taking aspirin assists in preventing stroke by)**		0.94	0.125
		Interpretation	Easy	Poor
9	**Anda berisiko tinggi untuk mendapat strok sekiranya** **(You are at greater risk of stroke if)**		1.00	0
		Interpretation	Easy	No discrimination
10	**Setelah mengalami strok kecil (TIA),** **(Once you have suffered a transient ischemic attack (TIA),**		0.88	0.25
		Interpretation	Easy	Acceptable
11	**Pembedahan kadang kala membantu untuk mengelakkan kejadian strok berulang melalui** **(Surgery can sometimes help to prevent another stroke by)**		0.94	0.125
		Interpretation	Easy	Poor
12	**Apakah kaedah rawatan yang terdapat bagi individu yang mengalami strok?** **(What method of treatment is available for people who have had a stroke?)**		0.88	0.25
		Interpretation	Easy	Acceptable
13	**Faktor risiko UTAMA penyebab strok ialah** **(The most important known risk factor for stroke is)**		0.88	0.25
		Interpretation	Easy	Acceptable
14	**Secara anggaran berapa ramaikah rakyat Malaysia mengalami serangan strok setiap tahun?** **(Approximately how many Australians are affected by stroke every year?)**		0.19	0.38
		Interpretation	Difficult	Good
15	**Sekiranya anda mengambil alcohol secara berlebihan, anda akan** **(If you drink alcohol excessively you are)**		0.69	0.63
		Interpretation	Moderate	Excellent
16	**Antara tanda-tanda berikut, manakah merupakan contoh masalah FIZIKAL akibat serangan strok?** **(Which of the following is an example of a physical disability caused by stroke?)**		0.88	0.25
		Interpretation	Easy	Acceptable
17	**Untuk mengurangkan risiko strok, anda perlu** **(To reduce the risk of stroke, you need to)**		1.00	0
		Interpretation	Easy	No discrimination
18	**Merokok 20 batang sehari meningkatkan risiko strok sebanyak** **(Smoking 20 cigarettes per day increases the risk of stroke by)**		0.50	0.75
		Interpretation	Moderate	Excellent
19	**Sekiranya seseorang mengalami strok, bilakah anda perlu menelefon ambulan?** **(If someone has a stroke, when should you call an ambulance?)**		1.00	0
		Interpretation	Easy	No discrimination.
20	**Rehabilitasi (pemulihan) strok dapat membantu seseorang yang** **(Rehabilitation can assist someone who has suffered)**		0.94	0.125
		Interpretation	Easy	Poor