Criterion Validity of the PHQ-9 (Malay Version) in a Primary Care Clinic in Malaysia

M S Sherina, PhD*, B Arroll, PhD, F Goodyear-Smith, MGP****

*Universiti Putra Malaysia, Malaysia, **University of Auckland, New Zealand

SUMMARY

Objective: This study was conducted to determine the validity of the 9-item Patient Health Questionnaire (PHQ-9) (Malay version) as a case-finding instrument for depression among women in a primary care clinic.

Methods: A cross sectional study was conducted in a primary care clinic in Malaysia. Consecutive adult women patients who attended the clinic during data collection were given self-administered questionnaires, which included the PHQ-9 (Malay version). Systematic weighted random sampling was used to select participants for Composite International Diagnostic Interviews (CIDI). The PHQ-9 was validated against the CIDI reference standard.

Results: The response rate was 87.5% for the questionnaire completion (895/1023), and 96.8% for the CIDI interviews (151/156). The prevalence of depression was 12.1% (based on PHQ-9 scores of 10 and above). The PHQ-9 had a sensitivity of 87% (95% confidence interval 71% to 95%), a specificity of 82% (74% to 88%), positive LR 4.8 (3.2 to 7.2) and negative LR 0.16 (0.06 to 0.40).

Conclusions: The Malay version of the PHQ-9 was found to be a valid and reliable case-finding instrument for depression in this study. Together with its brevity, it is a suitable case-finding instrument to be used in Malaysian primary care clinics.

KEY WORDS:
Validation, Depression, Primary Care, Women, Malaysia

INTRODUCTION

The Patient Health Questionnaire-9 (PHQ-9) is a self-report measure, consisting of nine questions based on the nine DSM-IV criteria for major depression. It is used to determine the presence or absence of depression. Its validity as a brief depression severity measure was first published in 2001 by Kroenke K, Spitzer RL and Williams JBW[1]. Its brevity coupled with its construct and criterion validity makes the PHQ-9 an attractive, dual-purpose instrument for making diagnoses and assessing severity of depressive disorders, as well as monitoring treatment.

The PHQ-9 has also been recommended as the best available screening / case-finding instrument for primary care based on its brevity, and ability to inform the clinicians on both depression severity and diagnostic criteria[2]. It has been found to be effective for the detection and monitoring of depression in diverse populations; such as in African-American, Chinese American, Latino and non-Hispanic white patient groups in primary care settings in the USA[3]. The PHQ-9 has been translated into many different languages. It has also been used in many different countries; such as Brazil, Canada, Denmark, Finland, France, Hong Kong, Italy, Korea, The Netherlands, Norway, Poland, Russia, Taiwan and USA[4]. There have been studies (not published) which validated the PHQ-9 in Malaysian settings, and one review article which recommended its use in Malaysia[5].

Malaysian national health and morbidity surveys have found an increasing trend of poor mental health status in the Malaysian community, which is consistently higher among women compared to men[6,7]. The Second National Health and Morbidity Survey (NHMS II) conducted from 1987 to 1996, found that the prevalence of psychiatric morbidity for people aged 16 years and above in Malaysia was 10.7% based on the 12-item General Health Questionnaire (GHQ-12), where women had 1.5 times higher rate of psychiatric morbidity compared to men[6]. While the Third National Health and Morbidity Survey (NHMS III) found that the prevalence of psychiatric morbidity had increased to 12.1% among Malaysian women based on the GHQ-28[7]. According to the Malaysian Burden of Disease Injury Survey conducted in 2004, major depression ranked third as the leading cause of disease burden in women, and tenth in men[8].

The primary goals of the 9th Malaysian Plan (2006-2010) were to reduce the burden of illness and enhance health care service delivery, where common mental health disorders such as depression and anxiety are a priority. Primary care was identified as the best setting for early detection and intervention[9].

This paper is part of a larger study which was funded under the 9th Malaysian Plan. The purpose of the study was to determine (1) the prevalence of depression and anxiety among women, and their associated factors, and (2) develop brief and valid case-finding instruments for depression and anxiety. This study investigated the validity of the Malay versions of the PHQ-9, the 7-item Generalized Anxiety Disorder (GAD-7), the two questions with help question (TQWHQ) and a single anxiety question with help question in a primary care clinic in Malaysia.

This article was accepted: 15 February 2012
Corresponding Author: Sherina Mohd Sidik, Department of Psychiatry, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia Email: sherina@putra.upm.edu.my

For the purpose of this paper, the validity results of the PHQ-9 (Malay version) as a case-finding instrument for depression are discussed. The types of validity determined here were translation validity and criterion-related validity. Findings on depression and anxiety, as well as the validity results of the GAD-7 and TQWHQ have been published elsewhere[10-13].

MATERIALS AND METHODS

This study was designed and analysed according to the STARD (STAndards for the Reporting of Diagnostic accuracy studies) statement[14]. A cross sectional study design was used and data were collected over a duration of eight weeks from 10 December 2009 to 30 January 2010. The PHQ-9 was validated in the Malay language against the Composite International Diagnostic Interview (CIDI) depression module.

Instruments

Patient Health Questionnaire (PHQ-9)

The PHQ-9 refers to symptoms experienced by patients during the two weeks prior to answering the questionnaire. After obtaining permission from the copy right holder, the PHQ-9 was translated following the guidelines for cross-cultural adaptation of self-report measures in this study[15]. The process included two independent forward translations of the original PHQ-9 into Malay, consensus between translators on the forward translation, back-translation by bilingual English teachers, and a review of the back-translation by an expert committee (content validity). See Appendix for the final version of the PHQ-9 (Malay version).

As a severity measure, the PHQ-9 scores range from 0 to 27, as each of the nine items was scored from 0 (not at all) to 3 (nearly every day). PHQ-9 scores of 5, 10, 15, and 20 represented mild, moderate, moderately severe, and severe depression, respectively[1]. In this study, a positive score was defined as score of 10 and above for PHQ-9, and participants with these scores were categorised as having depression.

General Health Questionnaire (GHQ-12)

The GHQ-12 is one of the most evaluated instruments in community primary care clinics, and has been found to be a reliable and valid instrument in different languages[16]. It is also widely used to assess psychiatric morbidity in Malaysia in hospital, primary care and community settings in both Malay and English languages[6,17-21]. Although it does not specifically detect depression or anxiety, it gives an overall estimate of the prevalence of poor mental health in the population tested. The GHQ-12 was used in this study to determine the construct validity of the PHQ-9.

The Composite International Diagnostic Interview (CIDI)

The CIDI was chosen as the reference standard for the PHQ-9 in this study due to its extensive and in-depth development as a psychiatric diagnostic instrument. It was developed to generate diagnoses based on the definitions and criteria of the World Health Organization International Classification of Disease (ICD), as this was especially important for cross-national comparative research[22]. The CIDI was validated against the Schedules for Clinical Assessment in Neuropsychiatry (SCAN), and found to be a highly valid assessment of non-psychotic disorders. Its concordance for ICD-10 diagnoses was found to be moderate to excellent (kappa=0.58-0.97)[23]. The CIDI was also found to be a reliable and practical instrument cross-culturally[24-26].

Selection of clinic

This study was carried out in a government primary care clinic in an urban district in Malaysia. An urban location was chosen due to findings of the national surveys that mental health problems are more prevalent in urban compared to rural settings[6,7]. The clinic was selected via simple random sampling from a list of government primary care clinics headed by a Family Medicine Specialist / Family Physician (FMS). This was a pre-determined criteria for the study due to safety issues, where participants diagnosed with depression and / or anxiety from the study were referred to the FMS for further consultation. The clinic provided outpatient, and maternal and child health care services by doctors, medical assistants and nurses.

Study Participants

Consecutive adult women patients who attended the clinic and fulfilled the selection criteria were approached by the Research Assistants (RAs). Inclusion criteria were all Malaysian citizens aged 18 years old and above. Exclusion criteria were patients who were acutely ill and needed to be attended to immediately, and those with communication problems. The RAs obtained written consent from each participant.

Study Design

Participants completed self-administered questionnaires (including the PHQ-9, GAD-7 and GHQ-12), and returned the questionnaires to the RAs. They did not receive any help in completing the questionnaires. Participants only took part once in this study, where those who had already answered the questionnaire were not sampled again if they returned to the clinic during the study duration. The RAs scored the PHQ-9 and GAD-7 of each questionnaire, and divided the participants into two main groups; (1) Normal scores (PHQ-9 < 10, and GAD-7 < 5) and (2) High scores (PHQ-9 ≥ 10 and / or GAD-7 ≥ 5).

For validation of the PHQ-9 and GAD-7, systematic weighted random sampling was conducted to select participants from both groups. One in ten participants with normal scores and one in two participants with high scores were selected for the CIDI. The weighted sample created a higher prevalence of depression and / or anxiety for the validation exercise (diagnostic interview with the CIDI), and enabled the sensitivity and specificity from the validation results to have a similar range of confidence intervals.

The validation exercise was conducted by only one investigator, who was the primary investigator (PI). The PI, a FMS, was blinded to the participants' scores and administered the CIDI in a consultation room which provided adequate privacy for the participants to disclose personal information if necessary. Questions were based on the CIDI-Auto software and participants' responses were keyed into the software which generated the diagnosis of the participants, based on the DSM-IV criteria. The interviews varied from 10 minutes to 90 minutes each, based on the participants' diagnosis. Flowchart of the study design is shown in Figure 1.

Table 1: Validity of the PHQ-9 compared with the CIDI as the reference standard for depression (N=146)

Patients screened Positive		Patients screened Negative		Sensitivity % (95% CI)	Specificity % (95% CI)	LR	
True Positive	False Positive	True Negative	False Negative			Positive (95% CI)	Negative (95% CI)
27	21	94	4	87 (71 to 95)	82 (74 to 88)	4.77 (3.17 to 7.19)	0.16 (0.06 to 0.40)

Data analysis
Data were entered into the statistical package for social sciences program (SPSS 16.0). They were carefully verified and checked again. Data from participants with known psychiatric illness and those who were on psychoactive drugs were excluded to ensure that the final analysed data did not include participants who were attending the clinic for pre-determined mental health reasons which could affect the results of the study.

Validity of the PHQ-9
The types of validity determined for the PHQ-9 in this study were: (i) translation validity, which included face validity and content validity, and (ii) criterion validity, which included concurrent validity and convergent validity.

Translation validity
Face validity was conducted by administering the Malay-translated version of the PHQ-9 in a pilot study to a group of respondents (n=80) to determine the "face values" of this instrument. The pilot test was conducted in a location not included in the study (which was another government primary care clinic), and included respondents with similar socio-economic background as in the proper study. The respondents were able to answer the questionnaires without any problems, although some clarifications on certain items were required. These items were noted down and some minor ammendments were made based on the comments of the respondents.

Content validity of the PHQ-9 was evaluated by a panel of experts who were academicians and clinicians from various fields. These experts were family medicine specialists, psychiatrists and public health specialists with clinical skills and knowledge, as well as research expertise for content validation of the PHQ-9. Based on their recommendations, some wording of the items was altered. Through this process, the content validity suitable to the purpose of the study was established.

Criterion validity
Concurrent validity of the PHQ-9 was determined by comparing the sensitivity, specificity and likelihood ratios (LRs) of the PHQ-9 with the CIDI (reference standard). The calculations were according to the calculator on the University of Toronto website (www.cebm.utoronto.ca) [27].

For convergent validity of the PHQ-9, Pearson's correlation coefficient was used to establish the PHQ-9 scores compared to the GHQ-12 scores.
To assess the effectiveness of the PHQ-9 as a diagnostic instrument for depression, the Receiver Operating Characteristic (ROC) curve and area under the ROC curve (AUC) were determined.

Ethics approval
Ethics approval was obtained from the (1) Ministry of Health, Malaysia, (2) Ethics Committee of the Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, (3) Clinical Research Centre of Malaysia, and (4) FMS-in-charge of the clinic.

RESULTS

Participants
Eight-hundred-and-ninety-five out of 1023 consecutive women patients fulfilling the selection criteria agreed to take part in the study (response rate of 87.5%). The age of the participants ranged from 18 to 81 years old (mean age 30.9±10.4). A majority of them were married (62.1%, 525/845).

Seven-hundred-and-thirty (82%) participants had normal scores in both PHQ-9 and GAD-7 questionnaires, while 165 (18%) participants had high scores in either one or both questionnaires. For the diagnostic interview with the CIDI, 75 participants were selected from the group with normal scores and 81 from the group with high scores.

Fifty questionnaires were excluded from the initially recruited patients for data analysis. This was due to missing data (n=30) and known psychiatric illness (n=20). Overall, 845 questionnaires (including 146 with CIDI data) were left for the final data analysis.

Using a cut-off point of 10 and above on the PHQ-9, 12.1% (102/845) participants were classified as having depression. Based on the CIDI (depression module), 31 out of the 146 participants interviewed were diagnosed as having depression.

Reliability and validity of the PHQ-9 (Malay version)
The Malay version of the PHQ-9 was found to have good internal reliability (Cronbach's alpha = 0.70)

For concurrent validity, the PHQ-9 had a sensitivity of 87%, specificity of 82%, a positive LR of 4.8, and a negative LR of 0.16 when compared to the CIDI as a reference standard (Table I).

As for convergent validity of the PHQ-9, the Pearson's correlation coefficient between the PHQ-9 and GHQ-12 was 0.61 ($p<0.001$). This indicated a positive association of moderate strength between the two instruments.

Based on the ROC curve (Figure 2), the AUC was found to have a value of 0.966, and the best (highest) sensitivity and specificity values were at the cut-off point of 10 and above for scores in the PHQ-9.

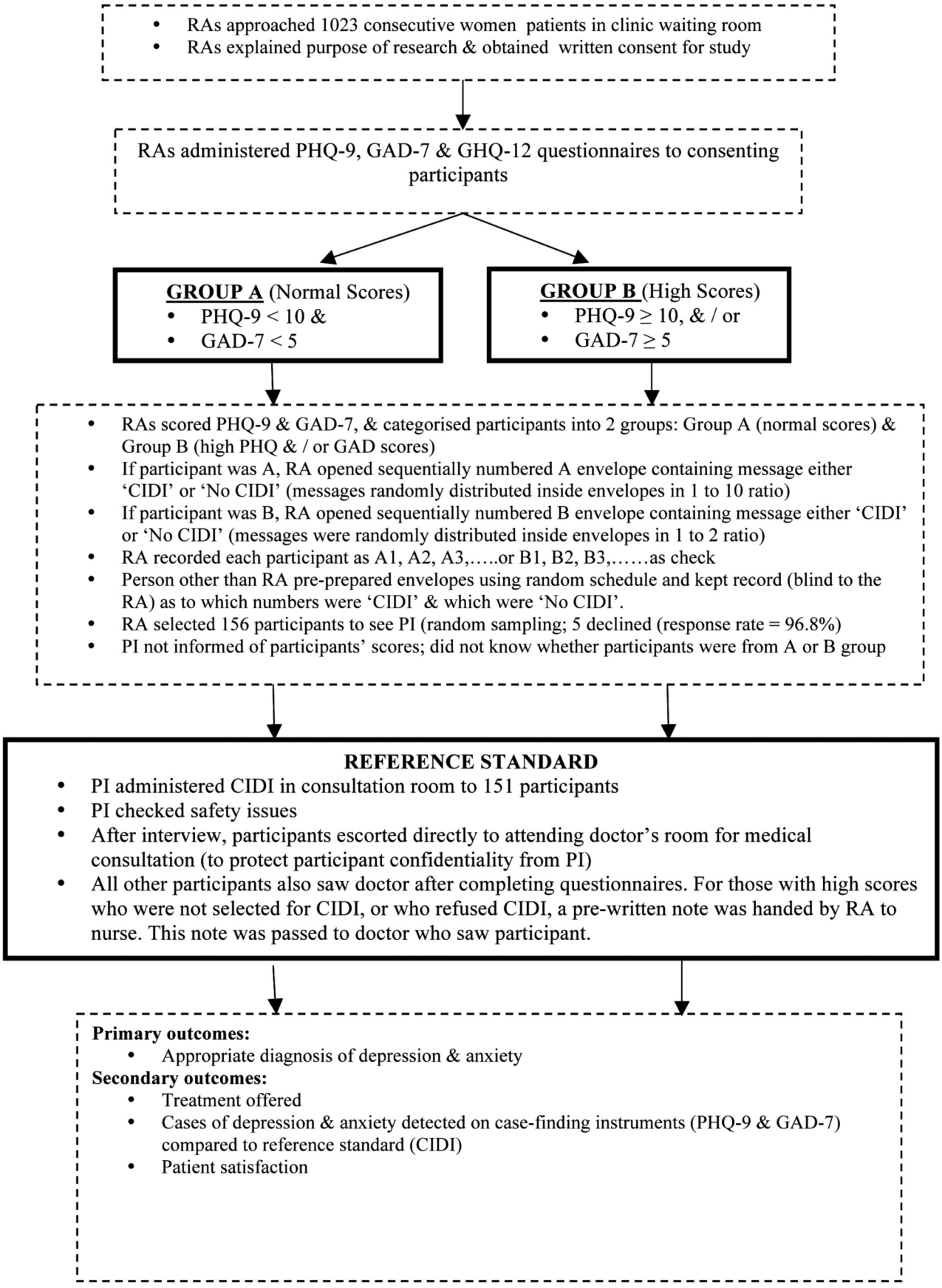

Fig. 1: Flowchart of study design.

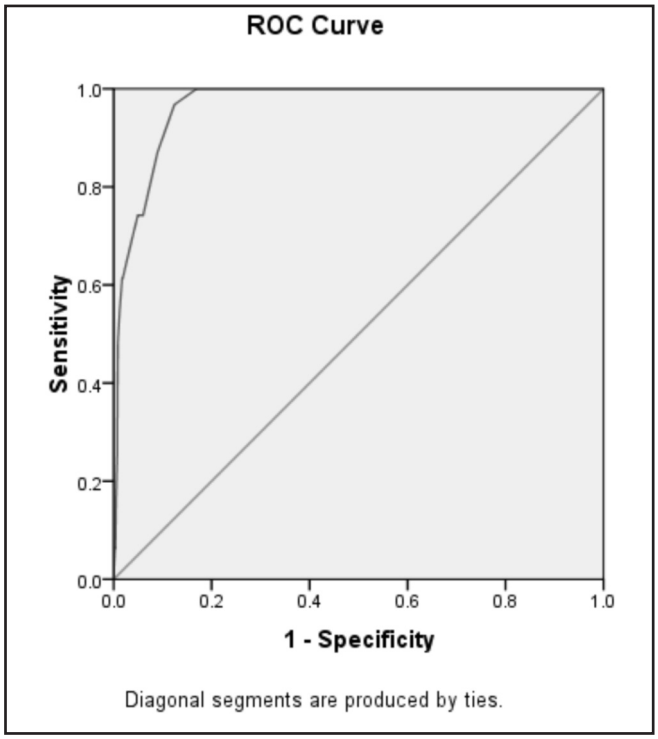

Fig. 2: Receiver Operating Characteristic (ROC) Curve of the PHQ-9 compared with the CIDI as the reference standard for depression (N=146).

DISCUSSION

Summary of main findings

The prevalence of depression in this study was 12.1%. The Malay version of the PHQ-9 was found to have excellent diagnostic performance as a case-finding instrument for depression as demonstrated by the ROC curve and AUC value. The PHQ-9 also had good sensitivity and specificity compared to the CIDI as the reference standard, which established good concurrent validity of the PHQ-9. Comparison of the PHQ-9 scores to those of the GHQ-12 established convergent validity of the PHQ-9.

Comparison with existing literature

This study found that the PHQ-9 had a reliability which was within the acceptable range. For a self-report instrument to be reliable, it is suggested that the Cronbach's alpha be at least 0.70 [28]. The internal consistency of the PHQ-9 in this study (Cronbach's alpha = 0.70) was almost similar to other studies from the United States (Cronbach's alpha = 0.79 – 0.89) [3,29], and another study from Thailand (Cronbach's alpha = 0.79)[30].

The sensitivity of the PHQ-9 at the cut-off value of 10 and above was 87% and the specificity was 82%. This showed a higher sensitivity, but a lower specificity compared to the findings of a recent validation study on the PHQ-9 in New Zealand[31]. Arroll et al found that at a similar cut off value of 10 and above, the PHQ-9 had a sensitivity of 74%, specificity of 91%, positive LR of 8.4 and negative LR of 0.28. However, the validity results of the PHQ-9 in this study is still acceptable, as the sensitivity of a screening instrument is considered to be good when the range is between 79% to 97%, and when the specificity is between 63% and 86% [32].

The positive LR of the PHQ-9 in this study was 4.8, and the negative LR was 0.16. This means that clinically in a similar primary care setting, a positive test with the PHQ-9 (cut-off score ≥10) would have almost a 30% chance of having depression and a negative test would have about a 2% chance of having depression[33]. These results show that the Malay version of the PHQ-9 is a suitable instrument for detecting depression among women in the primary care setting.

There was also satisfactory correlation between the PHQ-9 and GHQ-12 in this study (r=0.61, p<0.001), which confirmed the convergent validity of the PHQ-9. In a validation study of the PHQ-9 in Thailand, its convergent validity was compared to the Hamilton Rating Scale for Depression (HAM-D) and the PHQ-9 was also found to have good correlation with the HAM-D (r=0.56, p<0.001) [30].

Strengths of this study

Scientific importance

This study investigated in depth the validity of the PHQ-9 as a case-finding instrument for depression in a primary care setting in Malaysia. The translation of the PHQ-9 from its original language (English) to Malay was conducted according to standard guidelines for cross-cultural adaption of self report measures. Feedback from a team of experts ensured its content validity. The validation of the PHQ-9 in terms of concurrent validity against the CIDI is another major strength of this study. As the PHQ-9 was validated in Malay, this will greatly enhance its usage in primary care clinics and community settings throughout Malaysia. This study also has strengths in relation to its big sample size with proper probability sampling. It had a high response rate, and was designed and analysed according to the STARD statement[14].

Clinical importance

The effectiveness of the PHQ-9 as a diagnostic instrument for depression was demonstrated by the ROC curve, where the AUC was 0.966. Based on ROC curves, the closer the AUC value is to 1 the better the overall diagnostic performance of the instrument, and an instrument with an AUC value of 1 is one that is perfectly accurate[34]. Therefore the PHQ-9 (Malay version) had an excellent overall diagnostic performance as a case-finding instrument for depression in this study. The best (highest) sensitivity and specificity values were at the cut-off point of 10 and above for scores in the PHQ-9. With the PHQ-9, both the presence of clinical depression and also its degree of severity can be assessed[1,2].

Limitations of this study

A limitation (necessitated by time and resource constraints) is that the study was conducted among women in one government funded primary care clinic in an urban community setting, where the participants were mostly of lower to middle income socio economic status and therefore cannot be ensured to represent the Malaysian women population as a whole.

Predictive values were not discussed in this study because these values were all artificially high and affected by the prevalence of the disease[35]. This was due to the weighted selection of participants for the interview with the reference standard (CIDI). While this selection was necessary to enable the sensitivity and specificity for the concurrent validity to

have similar range of confidence intervals, it created an artificially high prevalence of disease.

CONCLUSION

The Malay version of the PHQ-9 has several potential advantages. It was found to be a valid and reliable case-finding instrument for detecting depression in primary care. It is shorter than other diagnostic instruments available in Malaysia for identifying depression. With the PHQ-9, the presence of clinical depression as well as its degree of severity can be assessed.

Competing interests

The authors declare that they have no competing interests.

ACKNOWLEDGEMENT

The study was funded under the Research University Grant Scheme (RUGS) from Universiti Putra Malaysia, Malaysia and the Postgraduate Research Student Support (PReSS) Accounts from the University of Auckland, New Zealand. Permission to conduct this study and publish this paper was obtained from the Ministry of Health, Malaysia and the Ethics Committee of the Faculty of Medicine and Health Sciences, Universiti Putra Malaysia.

REFERENCES

1. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. Validity of a brief depression severity measure. J Gen Intern Med 2001; 16: 606-13.
2. Nease DEJ, Maloin JM. Depression screening: a practical strategy. J Fam Pract 2003; 52(2): 118-24.
3. Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL. Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. J Gen Intern Med 2006; 21: 547-52.
4. List of questionnaires and translations: the PHQ-9. Lyon: MAPI Institute. [cited 2007, December 23]. Available from http://www.mapi-institute.com.
5. Zainab AM, Pereira XV. Depression in primary care. Part 1: screening and diagnosis. Malaysian Fam Physician 2007; 2(3).
6. Institute of Public Health. The Second National Health and Morbidity Survey 1996 (NHMS II). Kuala Lumpur: Ministry of Health Malaysia, 1999.
7. Institute of Public Health. The Third National Health and Morbidity Survey (NHMS III) 2006. Kuala Lumpur: Ministry of Health Malaysia, 2008.
8. Institute of Public Health. Malaysian burden of disease and injury study. Health prioritization: burden of disease approach. Kuala Lumpur: Ministry of Health Malaysia, 2004.
9. Economy Planning Unit Malaysia. Ninth Malaysia Plan 2006-2010. Putrajaya: Prime Minister's Department, 2006.
10. Sherina MS, Arroll B, Goodyear-Smith F. Prevalence of depression among women attending a primary urban care clinic in Malaysia. Singapore Med J 2012; 53(7): 1-6.
11. Sherina MS, Arroll B, Goodyear-Smith F. Validation of the GAD-7 (Malay version) among women attending a primary care clinic in Malaysia. J Primary Health Care 2012; 4(1): 5-11.
12. Sherina MS, Arroll B, Goodyear-Smith F, Azhar MZ. Screening for depression with a brief questionnaire in a primary care setting: Validation of the two questions with help question (Malay version). Intern J Psychiatry Med 2011; 41(2): 143-54.
13. Sherina MS, Arroll B, Goodyear-Smith F. The prevalence of anxiety among women attending a primary care clinic in Malaysia. Br J Gen Pract 2011; DOI:10.3399/bjgp.
14. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Br Med J 2003; 326: 41-4.
15. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000; 25(24): 3186-91.
16. Goldberg DP, Gater R, Sartorius N et al. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. Psychol Med 1997; 27: 191-7.
17. Quek KF, Low WY, Razack AH, Loh CS. Reliability and validity of the General Health Questionnaire (GHQ-12) among urological patients: a Malaysian study. Psychiatry Clin Neurosci 2001; 55: 509-13.
18. Jefferelli SB, Kong BH, Sherina MS, Azhar MZ. Mental health status of academic staff in a Malaysian medical school. Malaysian J Psychiatry 2003; 11(2): 12-7.
19. Jefferelli SB, Sherina MS, Nor Afiah MZ, Azlini O, Azhar MZ. The mental health status of government health care staff in Sepang district, Selangor. Malaysian J Psychiatry 2004; 12(1): 20-4.
20. Sherina MS, Rampal L, Kaneson N. Psychological stress among undergraduate medical students. Med J Malaysia 2004; 59(2): 207-11.
21. Noraini MN. Malaysian women's state of well-being: empirical validation of a conceptual model. J Soc Psychol 2006; 146(1): 95-115.
22. Kessler RC, Utsun TB. The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). Int J Methods in Psychiatr Res 2004; 13(2): 93-121.
23. Jordanova V, Wickramesinghe C, Gerada C, Prince M. Validation of two survey diagnostic interviews among primary care attendees: a comparison of CIS-R and CIDI with SCAN ICD-10 diagnostic categories. Psychol Med 2004; 34: 1013-24.
24. Wittchen HU, Robins LN, Cottler LB, Sartorius N, Burke JD, Regier D. Cross-cultural feasibility, reliability and sources of variance of the composite international diagnostic interview (CIDI). Br J Psychiatry 1991; 159: 645-53.
25. Wittchen HU. Reliability and validity studies of the WHO composite international diagnostic interview (CIDI): a critical review J Psychiatry Res 1994; 28(1): 57-84.
26. The International Consortium in Psychiatric Epidemiology. Harvard: ICPE. [cited 2008, May 20]. Available from http://www.hcp.med.harvard.edu/icpe/
27. Centre for evidence-based medicine [Internet]. Mount Sinai Hospital: CEBM. [cited 2009 June 20]. Available from: www.cebm.utoronto.ca
28. Dawson B, Trapp RG, editors. Basic and clinical biostatistics (3rd edition). United States: McGraw-Hill, 2001.
29. Lee PW, Schulberg HC, Raue PJ, Kroenke K. Concordance between the PHQ-9 and the HSCL-20 in depressed primary care patients. J Affective Disord 2007; 99(1-3): 139-45.
30. Lotrakul M, Sumrithe S, Saipanish R. Reliability and validity of the Thai version of the PHQ-9. BMC Psychiatry 2008; 8(46).
31. Arroll B, Goodyear-Smith F, Crengle S et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. Ann Fam Med 2010; 8(4): 348-53.
32. Robins LN, Wing J, Wittchen HU et al. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Arch Gen Psychiatry 1998; 45(12): 1069-77.
33. McGee S. Simplifying likelihood ratios. J Gen Intern Med 2002; 17: 647-50.
34. Park SH, Goo JM, Jo CH. Receiver operating characteristics (ROC) curve: Practical review for radiologists. Korean J Radiol 2004; 5: 11-8.
35. Deeks JJ, Altman DG. Diagnostics tests 4: likelihood ratios. Br Med J 2004; 329: 168-9.

APPENDIX: PATIENT HEALTH QUESTIONNAIRE (PHQ-9) (MALAY VERSION)

Dalam tempoh 2 minggu yang lepas, berapa kerapkali anda terganggu oleh masalah berikut? / *Over the last 2 weeks, how often have you been bothered by any of the following problems?*

No.	Questions	Coding
Q1	Sedikit minat atau keseronokan dalam melakukan kerja-kerja. *Little interest or pleasure in doing things.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3
Q2	Merasa murung, sedih atau tiada harapan. *Feeling down, depressed or hopeless.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3
Q3	Masalah hendak tidur / semasa tidur, tidur terlalu banyak. *Trouble falling / staying asleep, sleeping too much.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3
Q4	Merasa letih atau kurang bertenaga. *Feeling tired or having little energy.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3
Q5	Kurang selera atau terlalu banyak makan. *Poor appetite or over eating.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3
Q6	Mempunyai perasaan buruk terhadap diri sendiri – ataupun merasa gagal terhadap diri sendiri ataupun menghampakan diri atau keluarga. *Feeling bad about yourself – or that you are a failure or have let yourself or your family down.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3
Q7	Masalah menumpukan perhatian terhadap perkara-perkara seperti membaca suratkhabar atau menonton televisyen. *Trouble concentrating on things, such as reading the newspaper or watching television.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3
Q8	Bergerak atau bercakap dengan terlalu lambat sehingga disedari oleh orang lain. Ataupun bertentangan – terlalu resah atau gelisah sehingga anda bergerak lebih dari biasa. *Moving or speaking so slowly that other people could have noticed. Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3
Q9	Berfikiran bahawa lebih elok jika anda telah mati atau ingin mencederakan diri anda dalam sesuatu cara. *Thoughts that you would be better off dead or of hurting yourself in some way.*	Tidak pernah sama sekali / *Not at all* 0 Beberapa hari / *Several days* 1 Lebih dari seminggu / *More than half the days* 2 Hampir setiap hari / *Nearly everyday* 3