ARTICLE ORIGINAL

Validation of the Tunisian version of the Patient Health Questionnaire -9 (PHQ- 9) for Depression screening

Validation de la version Tunisienne du Patient Health Questionnaire -9 (PHQ- 9) pour le dépistage de la dépression

Rabaa Jomli[1], Houda Belhaj[2], Uta Ouali[1], Yosra Zgueb[1], Hend Jemli[1], Fethi Nacef[1]

1- Hôpital Razi / Université Tunis Elmanar/Faculté de médecine de Tunis
2- Hôpital Pierre Janet le havre France / Université du Havre

RÉSUMÉ

Introduction : L'utilisation de questionnaires validés et fiables est important pour le dépistage précoce de la dépression. Le PHQ-9 est une des échelles les plus recommandées à cause de sa simplicité et sa fiabilité dans le diagnostic positif de la dépression ainsi que sa sévérité.

Objectifs : valider la version Tunisienne du PHQ-9 pour le dépistage de la dépression

Méthodes : Notre étude était transversale, descriptive et analytique. Elle a porté sur un échantillon de 134 personnes tirés au sort dans la population du grand Tunis. Nous avons procédé à l'étude de la validation en se basant sur les tests suivants : l'étude de la validité d'apparence, la fidélité test-retest, l'étude de la cohérence interne, l'étude de la validité externe et nous avons évalué la prévalence dans l'échantillon de l'étude.

Résultats : Dans l'étude de la fidélité, le coefficient de corrélation intraclasse ICC était de 0,97 montrant une très bonne concordance avec un coefficient Kappa de Cohen à 0,602. La cohérence interne était évaluée par le coefficient alpha(0,84). Dans la validité externe, le coefficient de corrélation de Pearson était de 0,947. Pour un seuil de dépression à 10, le coefficient de Spearman était à 0,810, la sensibilité était de 86,2 et la spécificité à 83,8. La VPP était de 0,6 et la VPN était de 0,9. Le score moyen du PHQ-9 était de 8,1+/- 0,5. La prévalence ponctuelle de l'épisode dépressif majeur(EDM) pour notre échantillon était de 13,4%.

Conclusion : Notre étude de validation a montré que la version Tunisienne de l'échelle PHQ-9 présentait des caractéristiques psychométriques fiables. Elle pourrait donc être utilisée comme un outil de dépistage et d'évaluation de la sévérité de la dépression.

Mots clés : Dépression, Validation, Questionnaire, PHQ-9

SUMMARY

Bakground: Depression is a major burden for the health-care system worldwide. It is important to use standardized and easily applicable tools for large-scale population screening to improve the early detection of depression. The PHQ-9 has been recommended as the best available screening and case-finding instrument based on its brevity and ability to inform the clinicians on both depression severity and diagnostic criteria.

Aims: To evaluate the reliability and the validity of the Tunisian version of the PHQ-9 in the general population in Tunisia; the Tunisian version of the Hospital anxiety depression scale (HAD) was used as the gold standard for major depression diagnosis.

Methods: We undertook a cross-sectional, descriptive and analytical study. A total of 134 participants were enrolled. The PHQ-9 was validated against the HAD reference standard.

Results: Test-Retest reliability was determined by intraclass correlation. This scale was stable over 2 weeks (ICC=0,97).The Tunisian version of the PHQ-9 has a good internal reliability (Cronbach's alpha = 0.84). As for criterion validity of the PHQ-9, the Pearson's correlation coefficient between the PHQ-9 and HAD was 0.94 and the Spearman's correlation coefficient was 0,81. A cut-off score of 10 had a sensitivity of 86,2 and a specificity of 83,8. The VPP was 0,6 and the VPN was 0,9. The prevalence of depression in ours ample was 13,4%. There were no statistical difference in gender, education and age for the prevalence of depression.

Conclusion: The Tunisan version of the PHQ-9 is a valid and reliable casefinding instrument for detecting depression.

Key words: depression; validation; questionnaire; PHQ-9

Correspondance
Rabiaa Zaibi Jomli
Hôpital Razi / Université Tunis Elmanar/Faculté de médecine de Tunis
Email : rabaa.jomli@yahoo.fr

BACKGROUND

Depression is a major burden for the health-care system worldwide. Metaanalyses reported a prevalence ranging from 2% to 15% of all types of depression and a 32 % prevalence of mental health conditions in general [1]. The World Health Organization ranks depression as the fourth leading cause of disability worldwide [2], and projects that by 2020, it will be the second leading cause. It is associated to diminished quality of life, medical morbidity, and mortality [3]. Thus, the detection of depression and the dissemination of treatment in the general population are very important to reduce the burden of the disease. Therefore, it is important to use standardized and easily applicable tools for large-scale population screening to improve the early detection of depression. This method was recommended by the canadian task force on preventive health care, the US preventive services task force and the UK national institute of clinical excellence [4]. Compared with other instruments for screening depression, the PHQ-9 has been recommended as the best available screening and case-finding instrument based on its brevity and ability to inform the clinicians on both depression severity and diagnostic criteria [5].

AIMS

To evaluate the reliability and the validity of the Tunisian version of the PHQ-9 in a sample of the general population in Tunisia; the Tunisian version of the Hospital anxiety depression scale (HAD) was used as the gold standard for major depression diagnosis. Ous second aim was to and define socoiodemographic characteristics of screened positive to depression participants.

METHODS

Subjects

We undertook a cross-sectional, descriptive and analytical study. We proceeded to a random draw respecting the distribution of the Tunisian population according to age groups, gender and level of education by referring to the official statistics for the year 2014 http://www.ins.tn/. Therefore, we enrolled 134 participants, which is considered satisfactory given that for validation studies, a participant number of at least 10 times the number of questionnaire items is required. Participants were selected according to age, gender and educational level. Indeed, statistical analysis showed significant associations between PHQ-9 and HAD-D.

Data collection was conducted in December 2016. The PHQ-9 was validated against the HAD reference standard. Participants were recruited from neighbors, friends, colleagues and residents in Grand Tunis. The inclusion criteria in our study were participants' informed consent, an over 20 years old age, physical and psychological ability to complete the questionnaire. We excluded patients who had a severe mental disorder, communication disorders and hearing impairment.

PHQ-9

The PHQ-9 is the 9-item depression module of the Patient Health Questionnaire measuring depressing symptoms and its severity during the previous two weeks. It is a self-report questionnaire composed of nine items based on the DSM-IV criteria for assessing symptoms of depression [6]. The total score of the PHQ-9 ranged from 0 to 27. Each item of the PHQ-9 ranges from 0 to 3 (0 = not at all, 1 = several days, 2 = more than half of the days, and 3 = nearly every day). According to the recommendations of Kroenke et al [7], the severity of depressive symptoms is evaluated by calculating the summary score (mild depression from 5 to 9, moderate depression from 10 to 14, moderately severe depression from 15 to 19 and ≥20 indicates severe depression). Major depression is diagnosed if 5 or more of the 9 depressive symptom criteria have been present for at least "more than half the days" in the past 2 weeks and 1 of the symptoms is a depressed mood or anhedonia.

HAD

The Hospital Anxiety and depression scale was developed in 1983 by Zigmond and Snaith to screen depressive and anxious disorders among psychiatric and medical patients. It consists of two subscales: an anxiety subscale (HAD-A) and adepression subscale (HAD-D), containing each seven intermingled items.

This self-assessment tool is therefore composed of 14 items scored on a 4-point Likert scale ranging from 0 to 3 (0=never, 3= very often).

We chose the HAD scale for its reliability and simplicity. Comparable with PHQ-9, which is adopted by APA

and incorporating DSM-IV depression criteria, HAD scale remains the most widely used by non-psychiatric physicians.

Statistical analysis:

The types of validity determined for the PHQ-9 in this study were: translation validity, internal reliability and criterion validity. Three bilingual psychiatrists translated, individually, the scale into Tunisian Arabic version. Local Tunisian Arabic differs considerably from Modern Standard Arabic in language and grammar, and we decided therefore to use the local dialect in order to assure the best possible comprehension.

The translations were then discussed in a group consisting of all psychiatrists. The best translation for each item of the scale was decided by consensus of the group. The final translated scale was back translated to English by a bilingual expert who doesn't know the original version. The back translated scale was compared with the original one. The translated scale was pretested to a group of 30 other psychiatrists. Descriptive data on frequency, percentage, mean and standard deviations were used for the sociodemographic characteristics and the PHQ-9 questionnaire scores. The psychometric properties of the instrument were evaluated. Test–retest reliability analysis was done with the intraclass correlation coefficients (ICC). The internal consistency of the scale was assessed by Cronbach's coefficient. The concurrent validity of the PHQ-9 was evaluated by calculating the Pearson and Spearman correlation coefficients between the PHQ-9 and the HAD scale. For the validity criteria, the sensitivity, the specificity and the predictive values were calculated. The receiver operating characteristic (ROC) curve analysis was conducted to determine suitable cut-off points in the general population.

The data were analyzed using the Statistical Package for the Social Sciences 19.0 version (SPSS 19.0). Continuous variables were examined using a *t*-test, and dichotomous variables were analyzed using the chi-square test for group comparisons. A result was considered statistically significant if $p< 0.05$.

RESULTS

Sociodemographic characteristics of the participants and scores on the PHQ-9:

Our sample was composed of 134 participants aged between 20 and 74 years. The mean age was 40.96 years with a standard deviation of 14.71 years. There were 65 females (48.5%) and 69 males (51.5%). The sex ratio M/F was 1.06. Thirty-four participants were illiterate (25.4%), 43 had a primary education level (32.1%), 39 had a high school level (29.1%) and 18 participants had a university level (13.4 %). Mean PHQ-9 score was 8.1 (SD 0.5) with extremes ranging from 0 to 27. The point prevalence of major depressive episode in our sample according to categorical method was 13.4% (n=18). According to PHQ-9, the most frequent clinical symptoms found of depression were: feeling depressed (77.6%, n=104) and feeling tired (77.6%, n=104). Suicidal ideation was present in 27.6% (n=37). The prevalence of different severity of depressive symptoms according to the suggested cutoff points of having minimal (scores of 0-4), mild (scores of 5-9), moderate (scores of 10-14), moderately severe (scores of 15-19) and severe depressive symptoms (scores of 20-27) were respectively 32.8% (n=44), 35.8% (n=48), 14.9% (n=20), 11.9% (n=16), and 4.6% (n=6). Prevalence of Major depressive episode was higher in women (15.4%) compared to men (11.6%). Women had also higher scores on the PHQ-9 (the mean score of PHQ-9 was 8.4 in women and 7.4 in men).. Prevalence of depression was almost four times higher in participants with primary education level compared to those who have a university level (20.9% vs 5.6%). (Table 1and 2)

Psychometric properties

During this phase, the original questionnaire was translated into Arabic Tunisian dialect taking into consideration the peculiarities of the Tunisian culture. This translation method keeps an inferential equivalence of the instrument. We have chosen to validate a version in Arabic dialect. Indeed, it is recommended to use a native language version whenever possible.

Table 1. Items of PHQ-9

	Not at all 0	Several days 1	More than half the days 2	Nearly every day 3
Item 1(Little interest or pleasure in doing things)	47 (35,1%)	68 (50,7%)	7 (5,2%)	12 (9,0%)
Item 2 (Feeling down, depressed, or hopeless)	30 (22,4%)	67 (50,0 %)	15 (11,2 %)	22 (16,4%)
Item 3 (Trouble falling/staying asleep, sleeping too much)	39 (29,1 %)	49 (36,6%)	11 (8,2%)	35 (26,1 %)
Item 4 (Feeling tired or having little energy)	30 (22,4%)	67 (50,0 %)	13 (9,7 %)	24 (17,9%)
Item 5 (Poor appetite or overeating)	76 (56,7%)	33 (24,6%)	10 (7,5%)	15 (11,2%)
Item 6 (Feeling bad about yourself, or that you are a failure, or have let yourself or your family down)	80 (59,7%)	33 (24,6 %)	12 (9,0%)	9 (6,7 %)
Item 7 (Trouble concentrating on things, such as reading the newspaper or watching TV)	47 (35,1%)	63 (47,0%)	12 (9,0%)	12 (9,0%)
Item 8 (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 more than usual)	75 (56,0%)	34 (25,4 %)	11 (8,2%)	14 (10,4%)
Item 9(Thoughts that you would be better off dead or of hurting yourself in some way)	97 (72,4%)	22 (16,4 %)	7 (5,2%)	8 (6,0%)

Table 2. depression's items of HAD

	Not at all 0	Sometimes 1	Very often 2	Nearly all the time 3
Item 2 (I still enjoy the things I used to enjoy)	39 (29,1%)	41 (30,6%)	33 (24,6%)	21 (15,7%)
Item 4 (I can laugh and see the funny side of things)	66 (49,3%)	37 (27,6%)	24 (17,9%)	7 (5,2%)
Item 6 (I feel cheerful)	13 (9,7%)	72 (53 ,7%)	35 (26,1%)	14 (10,4%)
Item 8 (I feel as if I am slowed down)	57 (42,5%)	32 (23,9%)	26 (19,4%)	19 (14,2%)
Item 10 (I have lost interest in my appearance)	40 (29,9%)	21 (15 ,7%)	31 (23,1%)	42 (31,3%)
Item 12 (I look forward with enjoyment to things)	79 (59,0%)	37 (27,6%)	12 (9,0%)	6 (4 ,5%)
Item14 (I can enjoy a good book or radio or TV program)	59 (44,0%)	43 (32,1%)	14 (10,4%)	18 (13,4%)

Reliability

The 2-week test-retest reliability of the Tunisian version of the PHQ-9 was good, with a correlation coefficient of 0.97. Concordance between items' responses of the PHQ-9 between T0 and T15 was measured with Cohen's coefficient. Two items had a bad concordance (Cohen's Kappa between 0.21 and 0.41): item 3 assessing sleep disorders and item 4 evaluating the fatigue and the lack of energy. Item 2 had a moderate concordance, and item 1 a moderate to a good concordance. Items 6 and 8 had a good agreement (Cohen's Kappa between 0.61 and 0.80).

Items 5, 7 and 9 evaluating respectively appetite disorders, impaired concentration and suicidal ideas had a very good concordance.

The internal consistency, as assessed by Cronbach's alpha, was 0.84 for the full scale. (Figure 1)

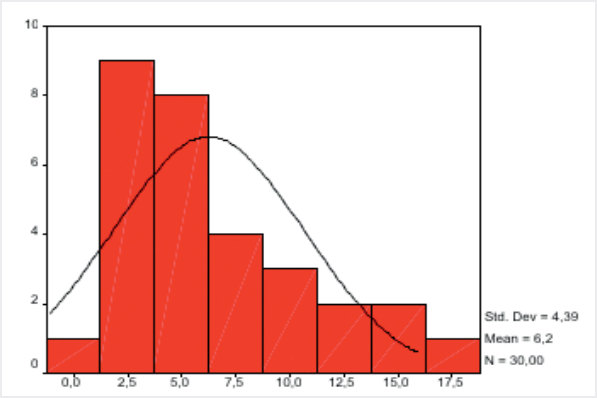

Figure 1. PHQ9 Score Global J

Criterion validity:

The two subscales correlated significantly with each other (Pearson's coefficient=0.94). The optimal cut-off point of 10 for the full scale yielded a sensitivity of 0.86 and a specificity of 0.83. The corresponding positive predictive value was 0.6 and the negative predictive value was 0.9. **(**Table 3)

Table 3: Concordance between items' responses of the PHQ-9

	Pearson's chi-squared	Cohen's Kappa
Item1 : Little interest or pleasure in doing things	0,00	0,6
Item2 : Feeling down, depressed, or hopeless	0,00	0,57
Item 3 : Trouble falling/staying asleep, sleeping too much	0,004	0,33
Item 4 : Feeling tired or having little energy	0,008	0,36
Item 5 : Poor appetite or overeating	0,00	0,83
Item 6 : Feeling bad about yourself, or that you are a failure, or have let yourself or your family down	0,00	0,75
Item 7 : Trouble concentrating on things, such as reading the newspaper or watching TV	0,00	0,94
Item8 : 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 more than usua	0,00	0,76
Item9 : Thoughts that you would be better off dead or of hurting yourself in some way	0,00	1,00

DISCUSSION

The PHQ-9, as a tool to screen assess severity of depressive symptoms, has gained growing appreciation for its brevity and tight match to the diagnostic criteria of the DSM-IV. In our study, we recruited a sample of 134 participants, selected randomly from the Tunisian general population. This study provided evidence on the psychometric properties of the PHQ-9, supporting its use as a reliable and valid measurement for depressive symptoms in Tunisian population.

Psychometric properties:

Reliability:

Good internal consistency and test-retest reliability of the PHQ-9 were demonstrated in our study. The internal reliability of the PHQ-9 was good; Cohen's Kappa coefficient showed for most items a medium to very good consistency. In the validation study of the primary version of the PHQ-9, the Kappa coefficient was 0.65 while in the study of Charlotte Hanlon et al, it was 0.21, which could be explained by choosing a cut off score of 5 [8]. This scale is stable over 2 weeks (test-retest reliability: 0.97), comparable to the test-retest reliability in previous studies. In the literature, the ICC varied from one study to another [9,10]. Thus, in the study of Xiaonan Yua et al, this coefficient was 0.76 [11], 0.92 in the Amharic version of PHQ-9 and 0.88 in the Dutch version. In our study the Cronbach's alpha was 0.84. To be reliable, it is desirable that a self-administrated questionnaire has a Cronbach's alpha greater than 0.70. This reflects a good internal consistency of our version of PHQ9 and its dimensionality. Indeed, the scale was developed so that all the items contribute to apprehend one dimension: depression. Our results are close to the US (α = 0.79-0.89) [12,13], the German (α = 0.87) and the Chinese (α = 0.86) PHQ-9 validation studies [14].

Criterion validity:

The two subscales correlated significantly with each other (Pearson's coefficient=0.84, Spearman's coefficient=0.81). These results were comparable to those published in literature [15,16]. In fact, in the Sri Lankan PHQ-9 validation study, the Pearson's coefficient was 0.87 and 0.61 in the Malaysian one. The optimal cut-off point of 10 for the full scale yielded a sensitivity of 0.86 and a specificity of

0.83. The corresponding positive predictive value was 0.6 and the negative predictive values was 0.9. In primary validation study of PHQ-9 conducted by Kroenke K and al, the threshold value was also 10 [17]. Literature data are in agreement with our results [18]. In other studies, the cut-off point was between 5 and 12 [17,18]; in Ethiopia, the cut point was lower (= 5) compared to other studies conducted in high-income countries [16], and it was high (12) in a study conducted in a hospital setting in patients with brain injury [19]. A screening test should have good sensitivity for better detection of affected individuals and a good negative predictive value to ensure that the excluded subjects have a low risk of being ill. Our study found a high sensitivity with a negative predictive value of the dialectal Arabic version of the PHQ-9. Our results do not deviate from the results reported in the literature, sensitivity varied between 28 and 100% and specificity raged between 88 and 98%. [20,21,22,23,11,24,25,26]. Several studies have found that the use of the categorical method for the detection of depression leads to a higher specificity of the test with the resulting lower sensitivity. Thus, screening for depression may be hampered by this calculation method [21]. However, this low sensitivity was observed only with the categorical method, but not with the threshold values.

For this reason, several authors recommend the use of a threshold approach rather than the algorithmic method [21] . The low sensitivity of the PHQ- 9 using the categorical method could also be explained by the choice of the gold standard or by cultural differences.

Prevalence of depression :

In our study, we found a point prevalence of depression of 13.4 %, according to the algorithmic method. Studies in other Tunisian populations showed that lifetime prevalence of depression was 8.6% in general population in 1995(Epidemiological study of depressive and psychotic disorders in the governorate of Ariana. Specialized thesis in psychiatry, defended on 6/10/1995 at the Faculty of Medicine of Tunis) and 26.4% in primary health care in 2013 [27]. In the general population, a multicenter epidemiological survey conducted in 2005 by the WHO; ‘ mental health in the general population : images and realities SMPG “ using the MINI , showed that the prevalence of depressive disorders in Tunis site was 37%[28] . The WHO study showed a considerably higher depression prevalence than we found in our study which could be due to different diagnostic criteria. Studies that have evaluated the prevalence of depression using the PHQ- 9 scale in the general population are fewer compared to those done in primary care settings. Although the prevalence of depression, in our study, was close to the national level, it was higher compared to international studies. For example, in Germany, the prevalence of depression was 3.8 % in 2006 and 2.7% in 2009 in the general population using the categorical approach of the PHQ- 9. Moreover, it seems that the prevalence of depression is higher when a threshold method is used rather than a categorical method. Indeed, it was 4.2 % in 2009 in Hong Kong for a cut-off of 9 and 8.1% in Germany for a cut-off of 10 [22,29]. In our study, using the global score, the minor and mild symptoms of depression accounted respective rates of 32.8 % (n = 44) and 35.8% (n = 48), followed by moderate ones with a rate of 14.9 % (n = 20). Moderately severe and severe forms accounted for respectively 11.9 %, (N = 16) and 4.6 % (N = 6). In a recent literature review on major depression in primary health care and in general population, authors found that minor to mild symptoms represent the most important symptoms compared to moderate to severe levels [30,14,22].

Prevalence of depression in our sample according to socio-demographic parameters:

It has been well documented that there is a gender difference in the prevalence of depression, with twice as many women compared with men [28 ,29, 14]. In our study, the prevalence of depression increased progressively with age to a peak of 25% (n= 4) between 60 and 79 years old. The difference was not significant. Studies have found that depressive symptoms are more common in elderly individuals, such as In the study of Kocalevent RD et al, the mean score of PHQ9 increases progressively with age from 2.3 in the 14 to 24 years age group to 4.4 among those aged over 75 years [24]. In our study, the prevalence of depression decreased with education level. Similarly, in other studies, using the PHQ-9, the prevalence of depression was higher among illiterate subjects (15.6%) and it diminished with grade level ; it was 5.6 % for a primary level of study, 4% for secondary level and 2.7% for university level [24,25].

Limitations of our study:

– The size of our population was small (n=134 participants). Larger populations were recruited in previous studies.

– The Arabic dialect of HAD is widely used in psychiatry and in several other disciplines but validation has been studied in literary Arabic.

What is already know on this topic:

– The prevalence of depression was 8.6% in general population in Tunisia in 1995

– There is no a validated questionnaire of depression in a Tunisian version

What this study adds:

– The PHQ-9 questionnaire in its Tunisian Arabic validated version shows a high sensitivity and specificity.

CONCLUSION

The PHQ-9 in its Tunisian Arabic validated version is a rapid and reliable screening instrument for depression, which also assesses symptom severity. This self-administered questionnaire can be used in psychiatry for diagnostic and prognosis purposes.

The Tunisian Arabic version of the PHQ-9 could improve the overall care of patients by facilitating early detection of depression.

Acknowledgements: On behalf of all authors, the corresponding author states that there is no conflict of interest.

REFERENCES

1. Singer S, Das-Munshi J, Brähler E. Prevalence of mental health conditions in cancer patients in acute care–a meta-analysis. Ann Oncol. 2010;21:925–30.
2. Kessler RC, Birnbaum HG, Shahly V, Bromet E, Hwang I, McLaughlin KA et al. Age differences in the prevalence and co-morbidity of DSM-IV major depressive episodes: results from the WHO World Mental Health Survey Initiative. Depress Anxiety. 2010, 27 (4): 351-364. 10.1002/da.20634.
3. Üstün TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ: Global burden of depressive disorders in the year 2000. Br J Psychiatry. 2004, 184: 386-392. 10.1192/bjp.184.5.386.
4. Macmillan HL, Patterson CJ, Wathen CN. The Canadian Task Force on Preventive Health Care. Screening for depression in primary care: recommendation statement from the Canadian Task Force on Preventive Health Care. CMAJ. 2005;172:33-5.
5. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. CMAJ. 2008; 178(8):997–1003.
6. Beaton D, Bombardier C, Guillemin F, Ferraz MB. Recommendations for the cross-cultural adaptation of health status measures. American Academy of Orthopaedic Surgeons.2002:1-9.
7. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a Brief Depression Severity Measure. J Gen Intern Med. 2001;16(9):606-13.
8. Hanlon C, Medhin G, Selamu M, Breuer E, Worku B, Hailemariam M. Validity of brief screening questionnaires to detect depression in primary care in Ethiopia. J Affect Disord. 2015;1(186):32-9.
9. Churchill GA. A paradigm for developing better measures of marketing constructs. Journal of marketing research. 1979;16:64-73.
10. Omoro SA, Fann JR, Weymuller EA, Macharia IM, Yueh B. Swahili translation and validation of the Patient Health Questionnaire-9 depression scale in the Kenyan head and neck cancer patient population. Int J Psychiatry Med. 2006;36:367-81.
11. Nicolaas PA, Vergouwe Y, King M, Nazareth I, Weze M, Moons K et al. The Patient Health Questionnaire-9 for detection of major depressive disorder in primary care: consequences of current thresholds in a crosssectional study. BMC family practice. 2010 ;11(1):11-98.
12. Streiner DL, Norman GR. Scaling responses in Health Measuremant Scales: a practical guide to their development and use. Oxford University Press. 1995;2:20–53.
13. 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(6):547-52.
14. Lee PW, Schulberg HC, Raue PJ, Kroenke K. Concordance between the PHQ-9 and the HSCL-20 in depressed primary care patients. J Affect Disord. 2007;99(13):139-45.
15. Kocalevent RD, Hinz A, Brähler E.Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry. 2013;35(5):551-5.
16. Hanwella R, Ekanayake S, de Silva VA. The Validity and Reliability of the Sinhala Translation of the Patient Health Questionnaire (PHQ-9) and PHQ-2 Screener. Depression Research and Treatment. 2014;(14):5.
17. Sherina MS, Arroll B, Goodyear-Smith F. Criterion validity of the PHQ-9 (Malay version) in a primary care clinic in Malaysia. Med J Malaysia. 2012;67(3):309-15.
18. Adewuya AO, Ola BA, Afolabi OO. Validity of the patient health questionnaire (PHQ-9) as a screening tool for depression amongst Nigerian university students. J Affect Disord. 2006;96:89–93.
19. Manea MS, Gilbody S, McMillan D. Optimal cut-off score for diagnosing

depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ. 2012;184(3):191-6.

20. Fann JR, Bombardier CH, Dikmen S, et al. Validity of the Patient Health Questionnaire-9 in assessing depression following traumatic brain injury. J Head Trauma Rehabil. 2005;20(6):501–11.

21. Kiely KM, Butterworth P. Validation of four measures of mental health against depression and generalized anxiety in a community based sample. Psychiatry Res. 2015;225(3):291-8.

22. Santos IS, Tavares BF, Munhoz TN, Almeida LS, Silva NT, Tams BD. Sensitivity and specificity of the Patient Health Questionnaire-9 (PHQ-9) among adults from the general population. Cad Saude Publica. 2013;(8):1533-43.

23. Yu X, Tam WW, Wong PT, Lam TH, Stewart SM. The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. Compr Psychiatry. 2012;53(1):95-102.

24. Volker D, Zijlstra-Vlasveld MC, Brouwers EP, Homans WA, Emons WH, Feltz-Cornelis CM. Validation of the Patient Health Questionnaire-9 for Major Depressive Disorder in the Occupational Health Setting. J Occup Rehabil. 2015;26(2):237-44.

25. Mathias K, Goicolea I ,Kermode M, Singh L, Shidhaye R, Sebastian MS. Cross-sectional study of depression and help-seeking in Uttarakhand,North India. BMJ. 2015;5(11):e008992.

26. Gelaye B, Williams MA, Lemma S, Deyessa N, Bahretibeb Y, Shibre T et al. Validity of the Patient Health Questionnaire-9 for depression screening and diagnosis in East Africa. Psychiatry Res. 2013;210(2):653-61.

27. Amamou B, Elkissi Y, Braham A, Souhaiel Bannour A, Ben Rejeb M, Ben Nasr S, Mtiraoui A, Ben Hadj Ali B. Prevalence and correlates of major depressive episodes in Sousse primary care setting: Assessment with Tunisian version of CIDI. La tunisie Medicale - 2013 ; 91 (04) : 234-239.

28. Cherif W, Elloumi H, Ateb A, Nacef F, Cheour M, Roelandt JL, Caria A. Représentations sociales de la santé mentale en Tunisie. La tunisie Medicale - 2012 ;90 (11) : 793 – 797.

29. Maske U, Busch M, Jacobi F, Beesdo-Baum K, Seiffert I, Wittchen H et al. Current major depressive syndrome measured with the Patient Health Questionnaire-9 (PHQ-9) and the Composite International Diagnostic Interview (CIDI): results from a cross-sectional population-based study of adults in Germany.BMC Psychiatry. 2015;10:15-77.

30. Hinz A, Mehnert A, Kocalevent RD, Brähler E, Forkmann T. Assessment of depression severity with the PHQ-9 in cancer patients and in the general population. BMC Psychiatry. 2016;(16):16-22.