HHS Public Access

Author manuscript

Asia Pac Psychiatry. Author manuscript; available in PMC 2018 September 01.

Published in final edited form as:

Asia Pac Psychiatry. 2017 September ; 9(3): . doi:10.1111/appy.12266.

The validity study of Patient Health Questionnaire-9 items for Internet screening in depression among Chinese University students

Na Du, **Kexin Yu**, **Yan Ye**, and **Shulin Chen**[1]

Department of Psychology, Zhejiang University, China

Abstract

Background—Internet intervention in Chinese University students would be a possible approach to overcome the gap between high rate of depression and high rates of underdiagnosis and undertreatment. As a popular measure of screening, the feasibility and user satisfaction of Patient Health Questionnaire-9 items for online program were tested.

Methods—The subjects were enrolled based on an email list from the students' office of a Chinese University, and 300 undergraduate students were randomly invited. Of which, 230 (76.7%) students were willing to participate in the study and completed the first test. After 2 weeks, a subsample of 150 (65.2%) subjects were randomly chosen to retake the test for the test–retest reliability. And 81 (35.2%) among the 230 subjects were randomly selected to undergo the Mini International Neuropsychiatric Interview (MINI) within 48 h. Among 150 subjects, 120 (52.2%) completed client satisfaction questionnaire about this online screening program.

Results—(1) The Cronbach's alpha was 0.80 and the test-retest reliability was 0.78; (2) the optimal cutoff score of 10 revealed a sensitivity of 0.74, specificity of 0.85, with an area under the curve of 0.897 (95% confidence interval: 0.823–0.970); (3) the mean duration of administration was 3.5 min; and (4) satisfaction with the online screening program was highly appreciated.

Conclusions—The results indicated potential value of the online screening program for further Internet-administrated programs of depression among Chinese University students.

Keywords

University students; Internet intervention; PHQ-9; feasibility; satisfaction

1. Introduction

According to the World Health Organization, about 350 million people are under depression, and a large proportion of them are below the age of 25 years (Kessler et al., 1993). In China,

[1]Corresponding author at: Department of Psychology, Zhejiang University, NO 148 Tianmushan Road, Xixi Campus of Zhejiang University, Hangzhou, Zhejiang 310028, China. Tel.: +86 571 88273337; fax: +86 571 88273326.

Contributors: Author Shulin Chen designed the study and instructed all the process of research. Author Na Du, Kexin Yu managed the literature searches and analyses. Author Na Du, Kexin Yu and Yan Ye conducted the experiment and collected the data. Author Na Du also undertook the statistical analysis and wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

the prevalence of depression among university students was 11.7% (Chen et al., 2013a), and depression was the major cause of functional disability in this population. However, mental health resources for managing depression are very limited in China, with only 1.46 psychiatrists per 100,000 population and 15 beds per 100,000 population for mental illnesses until 2011 (Qian, 2012). To overcome the gap between the high prevalence and the high rates of underdiagnosis and undertreatment for depression among Chinese University students is a big challenge.

The remarkable spread of Internet in the past 20 years in China makes its impact on the practice of mental health increasingly unignorable (Styra, 2004). In case of university students, who have developed trust and reliance on the Internet, web-based surveys to collect data on health risk behavior are feasible and cost-effective (Pealer et al., 2001). Meanwhile, students with depressive symptoms use the Internet much more often, which indicates that the students with depression are more likely to look for help from the Internet (Kotikalapudi et al., 2012). Hence, the Internet intervention program would be an efficacious approach for diminishing the gap among Chinese University students between the high rate of prevalence of depression and low rates of diagnosis and treatment. Groups of researchers have established numerous Internet intervention programs for depression such as MoodGYM, beating the blues, E-coach, Good days ahead, etc (Eells et al., 2014). A recent systematic review has indicated that mental health apps have the potential to be effective and may significantly improve treatment accessibility (Donker et al., 2013). Practical knowledge from these programs can be applied in similar Internet intervention programs in Chinese University students with depression. The Internet programs have advantages such as accessibility to health information, highly degree of anonymity, prompt feedback, cost-effectiveness and high capacity (Houston et al., 2001).

The Patient Health Questionnaire–9 (PHQ-9) was developed as a self-reported measure of depression, and it plays an important role in screening, assessment of changes, and self-management. With only nine items, the PHQ-9 meets five practical considerations, which include brevity, self-administration, multipurpose, present in the public domain, and easy to score (Kroenke, 2012a). Meanwhile, several studies have shown that the Chinese version of PHQ-9 is a valid and reliable tool for screening depression (Chen et al., 2013a, Yu et al., 2012, Wang et al., 2014, Liu et al., 2011). Evidence of feasibility of the PHQ-9 is also reported among Chinese University students (Zhang et al., 2013).

The objective of this study was to test the feasibility and user satisfaction of the online PHQ-9 screening program in Chinese University students. This is a preliminary step for the potential systematic Internet intervention of depression among Chinese University students.

2. Methods

2.1 Ethics statement

This study was performed in accordance with the Declaration of Helsinki and was conducted after getting approval from the Human Study Committee of Zhejiang University. Written informed consent was obtained from all study subjects before their enrollment.

2.2 Subjects

Subjects were enrolled from the undergraduate students of Zhejiang University. The inclusion criteria of the subjects were as follows: (a) age >18 years, (b) capable of independent communication, which includes the ability to listen, speak, read, and write, and (c) knowledge of basic computer skills, which include the ability to surf the Internet.

2.3 Procedures

Based on an email list from the students' office, 300 undergraduate students who have spent 1-3 years in the university were invited to participate in the online PHQ-9 screening program. Finally, 230 (76.7%) subjects were willing to participate in the study and completed the first test. After 2 weeks, a subsample of 150 (65.2%) subjects were randomly chosen to retake the test for the test–retest reliability.

After completing the online PHQ-9 screening program, 81 (35.2%) subjects from these 230 subjects were randomly selected to undergo the face-to-face MINI within 48 h. The MINI was administered by two research staff with experience in surveys using MINI and were blind to the PHQ-9 scores.

All the subjects were invited to complete client satisfaction questionnaires about this online PHQ-9 screening program, and 120 (52.2%) subjects responded to the questionnaires at the end.

2.4 Measures

2.4.1 PHQ-9—The PHQ-9 was developed as a self-reported, diagnostic screening tool for assessing and monitoring the severity of depression. It comprises nine diagnostic symptom criteria based on the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) major depressive disorder (MDD) (Kroenke et al., 2001). For each item, there are four response options ranging from “0 =Not at all” to “3 = Nearly every day” with total scores ranging from zero to 27. A major depressive episode is diagnosed in two ways using the PHQ-9: it is diagnosed using a diagnostic algorithm and a summary score (Spitzer et al., 1999). In this study, the Chinese version of the PHQ-9 and the diagnostic algorithm was used.

2.4.2 Mini International Neuropsychiatric Interview—The MINI is the gold standard among the various diagnostic tools. This tool is a short diagnostic structured interview for the major Axis I psychiatric disorders in DSM-IV (Lecrubier et al., 1997). The MINI translations were produced to ensure concordance with existing translations (Boudrot et al., 2009), and the Chinese version showed high sensitivity (Se: 0.94) and specificity (Sp: 0.86) for diagnosing a current major depressive episode, indicating its reliability and validity (Si, 2009). In this study, depression was diagnosed using the Chinese depression modules of the MINI.

2.4.3 Client satisfaction questionnaire—The client satisfaction questionnaire was compiled by researchers. It is a seven-point Likert-type self-reported instrument that assesses people's satisfaction toward the online PHQ-9 screening program in three

dimensions (convenience, comprehensiveness, and willingness). The overall satisfaction is also measured in this manner. In this questionnaire, subjects must react on a scale ranging from *strongly disagree* to *strongly agree*.

2.5 Statistical analysis

The data were analyzed using the Statistical Package for the Social Sciences (SPSS) Version 19.0. In case of the reliability analysis, the internal consistency and the test–retest consistency were assessed using Cronbach's alpha coefficient and Spearman's correlation coefficient, respectively.

In case of the validity criteria, the Se, the Sp, the predictive values, and the likelihood ratios of various intervals were calculated. The receiver operating characteristic (ROC) curve analysis was performed to determine the suitable cutoff points in the undergraduate students using Internet-administered PHQ-9.

Feasibility measures of interest including time of administration and difficulties in administration were addressed and analyzed. The scores of the satisfaction questionnaire in three dimensions were also addressed and analyzed.

3. Results

3.1 Demographic characteristics

The mean age of the subjects was 19.93 years (standard deviation, =2.58 years; range, 18-25 years). There were 130 (56.5%) men and 100 (43.5%) women.

Demographic data are presented in Table 1. The ratio of men to women was not significantly different (χ^2=0.101, P=0.951). The analysis of variance (ANOVA) showed no significant difference in mean age between men (F=1.230, P=0.296) and women (F=0.226, P=0.767). Meanwhile, as expected, the ANOVA showed significant differences in PHQ-9 total scores between men (F=38.403, P<0.01) and women (F=54.361, P<0.01).

3.2 Feasibility

3.2.1 Reliability—The internal consistency of Internet-administered PHQ-9 was 0.8, as established by Cronbach's alpha coefficient. The 2-week test–retest reliability coefficient of Internet-administered PHQ-9 was 0.78. The correlations between nine items of the PHQ-9 and the total scores ranged from 0.54 to 0.69 (Table 2).

3.2.2 Validity—The performance of the PHQ-9 against the diagnosis of major depression using the MINI as a standard criterion was examined. The sensitivity, specificity, and likelihood ratio of different cutoff scores were demonstrated (Table 3). At the cutoff score of the PHQ-9 of ≥10, the sensitivity was 0.74 and the specificity was 0.85. The positive likelihood ratio was 5.08 at this cutoff point. ROC curve analysis supports the criterion validity of the PHQ-9 in successfully differentiating between patients with and without major depression [area under the curve (AUC) =0.897, standard error=0.038, 95% confidence interval=0.823–0.970] (Figure 1).

3.2.3 Time-cost—The duration of administration of the online PHQ-9 screening program for university students ranged from 1 to 7 min, and the mean duration was 3.5 min. No major difficulties in administration were reported; all the students were able to complete the PHQ-9 and none of them objected to the contents of the items.

3.3 Total satisfaction score

With regard to the seven-point Likert-type satisfaction questionnaire, the dimension concerning convenience was judged most positively (M=5.71), and willingness was mostly judged neutral (M=5.43). However, the dimension concerning comprehensiveness was judged little negatively (M=4.86). Above all, the overall satisfaction was found to be 5.31. It means that though some limitations exist, university students' satisfaction toward the online PHQ-9 screening program is rather highly appreciated.

4. Discussion

With high feasibility and satisfaction, PHQ-9 is a powerful and potential online screening tool for depression. The overall Cronbach's alpha of the Chinese version of the Internet-administrated PHQ-9 was 0.8 in this sample and the 2-week test–retest reliability coefficient was 0.78. These reliability and validity are acceptable and comparable to the other studies on the Chinese PHQ-9 version, in which the reliability coefficients and test–retest reliability coefficients ranged from 0.80 to 0.89 and from 0.76 to 0.89, respectively (Wang et al., 2014, Yu et al., 2012, Chen et al., 2013b, Liu et al., 2011, Fann et al., 2009, Pinto-Meza et al., 2005, Zhang et al., 2013).

The screening properties of the Internet-administered PHQ-9 were supported by excellent AUC values. The sensitivity and specificity of the PHQ-9 of the present study did not differ from most studies that used PHQ-9 in a primary care setting, in spite of the different sample population of the present study (Gilbody et al., 2007). Kroenke (2012b) demonstrated that the optimal cutoff score of PHQ-9 may vary depending on the setting, and it was found to have acceptable diagnostic properties for detecting MDD for cutoff scores between 8 and 11. In the present study, the sensitivity and specificity of the PHQ-9 reached a satisfactory balance when the cutoff score was 10, which is lower than the value from the study by Zhang et al. (2013) who used paper-based PHQ-9 to screen depression among Chinese University students. The reason may be that in the present study, as the PHQ-9 was administered on the Internet, the subjects might not realize their depressive symptoms comprehensively, and some of them were even reluctant to admit the existence of depressive symptom to themselves. Above all, further studies on the cutoff scores of the Internet-administered PHQ-9 in university population are necessary to test this possibility.

Internet-based therapies have been proven to have many advantages over traditional ones. In contrast to people's general disinclination to consult physicians, the number of subjects enrolled in study accounted for 76.7% of the total students invited, which suggests high engagement of users. Meanwhile, the Internet-administered PHQ-9 takes shorter time and lesser effort to complete than the paper version (mean duration of administration =7.5 min) (Chen et al., 2010). It can be concluded that the subjects were mostly satisfied with this approach as a tool for screening depression, as shown by the results of the satisfaction

questionnaire. Online PHQ-9 also eliminates the barriers of distance, reduces stigma, and provides a cost-efficient platform. Given these advantages, it can be pictured that a combination of PHQ-9 and Internet has underlying applications.

First, PHQ-9 is a depression module which scores each of the nine DSM-IV criteria of depression, and which helps to increase the emotional self-awareness of its user. Second, the Internet-based PHQ-9 screening can also be adopted as a tool for the self-management of depression. Conducting tests at fixed intervals can help in a measure-based intervention. Third, current health app developers are driven mostly by economic motivation instead of research motivation (Donker et al., 2013), and hence the test of internet-administrated PHQ-9 is a preliminary step toward an evidence-based intervention program.

Limitations of the present study and of Internet screening must be acknowledged. First, as the sample size was small and limited within the scope of the university students, the results cannot be applied to other population. Second, only MINI was used as the gold standard diagnoses; it will be more convincing if more indicators are adopted. Finally, as the study selectively excludes subjects without access to the Internet the results may be biased. Future studies on online PHQ-9 screening program can explore the feasibility of the PHQ-9 as a self-management of depression or measurement-based intervention tool.

Overall, the results of the present study provide support for the online PHQ-9 screening program to assess and monitor the severity of depression among Chinese University students. It also indicates the PHQ-9 is a possible way to overcome the gap in the high rate of prevalence of depression and low rates of diagnosis and treatment.

Acknowledgments

The project described was supported by grants: R01TW008699 from the Fogarty International Centre of National Institutes of Health, USA; R01MH100298 from the National Institutes of Mental Health, USA; and the Program for New Century Excellent Talents in Universities of China.

We would like to acknowledge Mr. Lei Ding, who provided assistance during the data collection process and Mr. Qi Wang for his technical support for the online screening program.

References

Boudrot A, D'uva F, Sheehan D, Lecrubier Y, Even C. Pmc66 Use of the Mini International Neuropsychiatric Interview (M.I.N.I.) – Version 6 – in an International Study. Value in Health. 2009; 12:A31–A32.

Chen L, Wang L, Qiu XH, Yang XX, Qiao ZX, Yang YJ, Liang Y. Depression among Chinese university students: prevalence and socio-demographic correlates. PLoS One. 2013a; 8:e58379. [PubMed: 23516468]

Chen S, Chiu H, Xu B, Ma Y, Jin T, Wu M, Conwell Y. Reliability and validity of the PHQ-9 for screening late-life depression in Chinese primary care. Int J Geriatr Psychiatry. 2010; 25:1127–33. [PubMed: 20029795]

Chen SL, Fang Y, Chiu HL, Fan HN, Jin T, Conwell Y. Validation of the nine-item Patient Health Questionnaire to screen for major depression in a Chinese primary care population. Asia-Pacific Psychiatry. 2013b; 5:61–68. [PubMed: 23857806]

Eells TD, Barrett MS, Wright JH, Thase M. Computer-Assisted Cognitive-Behavior Therapy for Depression. Psychotherapy. 2014; 51:191–197. [PubMed: 24059735]

Fann JR, Berry DL, Wolpin S, Austin-seymour M, Bush N, Halpenny B, Lober WB, Mccorkle R. Depression screening using the Patient Health Questionnaire-9 administered on a touch screen computer. Psychooncology. 2009; 18:14–22. [PubMed: 18457335]

Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the patient health questionnaire (PHQ): A diagnostic meta-analysis. Journal of General Internal Medicine. 2007; 22:1596–1602. [PubMed: 17874169]

Houston TK, Cooper LA, Vu HT, Kahn J, Toser J, Ford DE. Screening the public for depression through the Internet. Psychiatric Services. 2001; 52:362–367. [PubMed: 11239106]

Kessler RC, Mcgonagle KA, Swartz M, Blazer DG, Nelson CB. Sex and Depression in the National Comorbidity Survey .1. Lifetime Prevalence, Chronicity and Recurrence. Journal of Affective Disorders. 1993; 29:85–96. [PubMed: 8300981]

Kotikalapudi R, Chellappan S, Montgomery F, Wunsch D, Lutzen K. Associating Internet Usage with Depressive Behavior Among College Students. Ieee Technology and Society Magazine. 2012; 31:73–80.

Kroenke K. Enhancing the clinical utility of depression screening. Canadian Medical Association Journal. 2012a; 184:281–282. [PubMed: 22231681]

Kroenke K. Enhancing the clinical utility of depression screening. CMAJ. 2012b; 184:281–2. [PubMed: 22231681]

Kroenke K, Spitzer RL, Williams JBW. The PHQ-9 - Validity of a brief depression severity measure. Journal of General Internal Medicine. 2001; 16:606–613. [PubMed: 11556941]

Lecrubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Sheehan KH, Janavs J, Dunbar GC. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. European Psychiatry. 1997; 12:224–231.

Liu SI, Yeh ZT, Huang HC, Sun FJ, Tjung JJ, Hwang LC, Shih YH, Yeh AWC. Validation of Patient Health Questionnaire for depression screening among primary care patients in Taiwan. Comprehensive Psychiatry. 2011; 52:96–101. [PubMed: 21111406]

Pealer LN, Weiler RM, Pigg RM, Miller D, Dorman SM. The feasibility of a Web-based surveillance system to collect health risk behavior data from college students. Health Education & Behavior. 2001; 28:547–559. [PubMed: 11575685]

Pinto-Meza A, Serrano-Blanco A, Penarrubia MT, Blanco E, Haro JM. Assessing depression in primary care with the PHQ-9: can it be carried out over the telephone? J Gen Intern Med. 2005; 20:738–42. [PubMed: 16050884]

Qian JW. Mental Health Care in China: Providing Services for Under-treated Patients. Journal of Mental Health Policy and Economics. 2012; 15:179–186. [PubMed: 23525836]

Si TM, Shu L, Dang WM, Su YA, Chen JX, Dong WT, et al. Zhang WH. Evaluation of the reliability and validity of Chinese version of the Mini-International Neuropsychiatric Interview in patients with mental disorders. Chinese Mental Health Journal. 2009; 23:493–497.

Spitzer RL, Kroenke K, Williams JBW, Primary PHQ. Validation and utility of a self-report version of PRIME-MD - The PHQ primary care study. Jama-Journal of the American Medical Association. 1999; 282:1737–1744.

Styra R. The Internet's impact on the practice of psychiatry. Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie. 2004; 49:5–11. [PubMed: 14763672]

Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, Zhang G, Zhou Q, Zhao M. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry. 2014; 36:539–44. [PubMed: 25023953]

Yu XN, Tam WWS, Wong PTK, Lam TH, Stewart SM. The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. Comprehensive Psychiatry. 2012; 53:95–102. [PubMed: 21193179]

Zhang YL, Liang W, Chen ZM, Zhang HM, Zhang JH, Weng XQ, Yang SC, Zhang L, Shen LJ, Zhang YL. Validity and reliability of Patient Health Questionnaire-9 and Patient Health Questionnaire-2 to screen for depression among college students in China. Asia Pac Psychiatry. 2013; 5:268–75. [PubMed: 24123859]

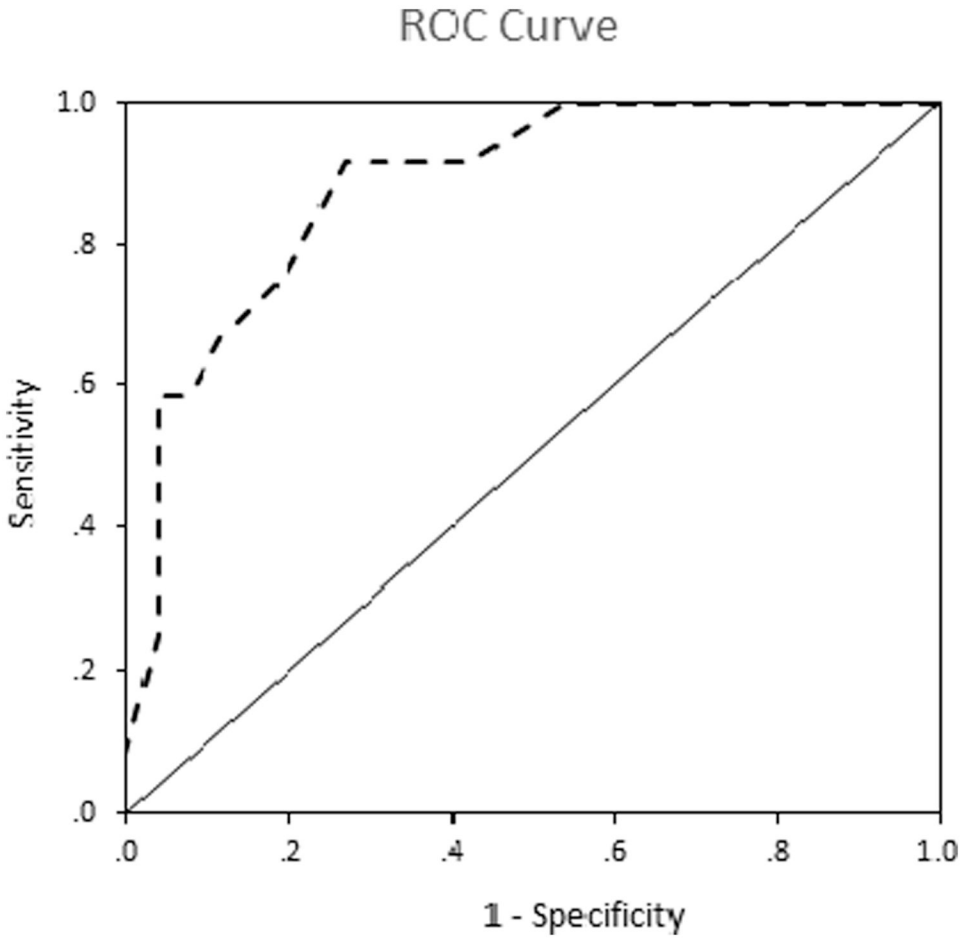

Fig.1.

ROC curve of the Chinses versions of the PHQ-9 versus the MINI. MINI=Mini International Neuropsychiatric Interview, PHQ-9=Patient Health Questionnaire, ROC=receiver operating characteristic

Table 1

Characteristics of subjects

	NON	MINOR	MAJOR	Statistics	*P*
	N=204	**N=17**	**N=9**		
Gender[a]					
Men	116	9	5		
Women	88	8	4	χ^2=0.101	0.951
Age[b]					
Men	20.1 (0.1)	20.0 (0.5)	19.2 (0.6)	F(2,127)=1.230	0.296
Women	19.8 (0.1)	19.8 (0.3)	20.3 (0.5)	F(2,97)=0.226	0.767
PHQ-9 Score					
Men	5.6 (0.3)	10.1 (1.1)	17.0 (1.5)	F(2,127)=38.403	*P*<0.01 *
Women	5.4 (0.3)	12.9 (0.3)	16.8 (1.4)	F(2,97)=54.361	*P*<0.01 *

[a] Chi-square test.

[b] Analysis of variance.

* Statistically significant.

MAJOR=major depression, MINOR=minor depression, NON=non-depression; PHQ-9=Patient Health Questionnaire

Table 2

PHQ-9 item level values and item-total correlations (n=230)

PHQ-9	Mean	SD	Item-total correlation
(1) Little interest or pleasure in doing things	0.91	0.67	0.63
(2) Feeling down, depressed, or hopeless	0.85	0.59	0.59
(3) Trouble falling or staying asleep, or sleeping too much	0.60	0.82	0.54
(4) Feeling tired or having little energy	0.86	0.71	0.69
(5) Poor appetite or overeating	0.67	0.82	0.63
(6) Feeling bad about yourself—or that you are a failure	0.77	0.78	0.66
(7) Trouble concentrating on things	1.07	0.90	0.61
(8) Moving or speaking so slowly that other people could have noticed	0.47	0.74	0.58
(9) Thoughts that you would be better off dead or of hurting yourself	0.18	0.48	0.55

PHQ-9=Patient Health Questionnaire, SD=standard deviation.

Table 3

Sensitivity, specificity, and likelihood ratios at various cutoff points of the PHQ-9 (n=81)

PHQ-9 score	Sensitivity	Specificity	LR+	LR-
≥5.00	0.95	0.53	2.03	0.10
≥6.00	0.95	0.65	2.67	0.08
≥7.00	0.95	0.69	3.09	0.08
≥8.00	0.89	0.74	3.47	0.14
≥9.00	0.89	0.79	4.27	0.13
≥10.00	0.74	0.85	5.08	0.31
≥11.00	0.63	0.92	7.83	0.40
≥12.00	0.53	0.94	8.16	0.51
≥13.00	0.47	0.94	7.34	0.56
≥14.00	0.42	0.97	13.05	0.60
≥15.00	0.37	0.97	11.42	0.65

LR+=positive likelihood ratio, LR-=negative likelihood ratio, PHQ-9=Patient Health Questionnaire.