Validation of the Patient Health Questionnaire-9 for Depression Screening Among Chinese Americans

Albert Yeung, M.D., Sc.D.[1,2], **Freddy Fung, B.A.**[1], **Shu-Ching Yu, MA**[1], **Sienna Vorono, B.A.**[1], **Mary Ly, MPH**[1], **Shirley Wu, B.S.**[1], and **Maurizio Fava, M.D.**[1]

[1]*Depression Clinical and Research Program, Massachusetts General Hospital*

[2]*South Cove Community Health Center*

Abstract

Objective—To validate the Chinese Bilingual version of the Patient Health Questionnaire (CB-PHQ-9) for screening for depression among Chinese Americans in primary care.

Method—Chinese American patients attending a primary care clinic were screened for major depressive disorder (MDD) using the CB-PHQ-9. All patients who scored 15 or higher on the CB-PHQ-9 were interviewed using the Chinese-Bilingual Structured Clinical Interview for DSM-IV, patient version (CB-SCID-I/P). Eight percent of patients who scored less than 15 on the CB-PHQ-9 were randomly selected and interviewed with the depression module of the CB-SCID-I/P.

Results—During the study, 1,940 unduplicated patients completed the CB-PHQ-9. The CB-PHQ-9 had high internal consistency (alpha=0.91), and the scores of each of its nine items had moderate to moderately high correlations (0.52-0.85) with the total scores. The sensitivity and specificity of the CB-PHQ-9 for recognizing MDD were found to be 81% and 98% respectively, and the area under the curve of the Receiver Operative Characteristic curve was estimated to be 0.97 (s.e. 0.01).

Conclusion—The CB-PHQ-9 is a valid and useful instrument for screening for MDD among Chinese Americans in primary care.

Keywords

Depression; Screening; Primary Care; Chinese Americans; CB-PHQ-9

Introduction

Major depressive disorder (MDD) is common among immigrant Chinese Americans, who tend to under-utilize mental health services (1) and seek help from primary care practitioners (PCPs) when they suffer from the disorder (2). Under-recognition of MDD in this population continues to be frequent despite efforts to increase awareness among primary care practitioners and among patients (3). It is therefore important for PCPs to be able to recognize when Asian Americans are depressed. Routine screening of MDD in primary care settings may be a possible remedy to resolve this public health challenge (4).

In an earlier study, our team at the Massachusetts General Hospital investigated the Chinese version of the 21-item Beck Depression Inventory (CBDI), and showed that it was a valid instrument for screening for depression among Chinese Americans (5). Yet, a considerable number of patients in the study complained about the length of the questionnaire and declined to participate, thereby limiting the usefulness of the CBDI as a screening instrument.

The PHQ-9 is the depression module of the Patient Health Questionnaire (PHQ), a criteria-based instrument for screening for mood, anxiety and other psychiatric disorders. The PHQ-9 consists of one item for each of the nine criterions upon which the diagnosis of DSM-IV major depressive disorders are based. In a validation study performed on 580 patients recruited from primary care and obstetrics-gynecology clinics, Kroenke et al. (8) used PHQ-9 ≥ 10 as the threshold and showed that the instrument had a sensitivity of 88% and a specificity of 88% for identifying MDD. He suggested that PHQ-9 scores of 5, 10, 15, and 20 represent mild, moderate, moderately severe, and severe depression respectively. Because of its brevity and its relatively high positive predictive value (6), the PHQ-9 has gained increased recognition as the preferred instrument for researching depression screenings in primary care settings (8-11).

To our knowledge, there has been no published validation study on using the PHQ-9 for screening for depression among Chinese Americans. Our team translated the PHQ-9 into Chinese, and back-translated it into English for comparison. The Chinese version was later finalized after several revisions with input from our bilingual clinicians. We added the Chinese translated version of the PHQ-9 to the original questionnaire to compile the Chinese Bilingual PHQ-9 (CB-PHQ-9)(appendix). In this study, we hypothesize that the CB-PHQ-9 is a valid instrument for MDD screening among immigrant Chinese Americans.

METHOD

Subjects

Subjects were Chinese-Americans, 18 years of age or older, who attended South Cove Community Health Center (South Cove) between September 15, 2004 to February 28, 2005. South Cove is an urban community health center in Boston. It serves low-income Asian immigrants who face financial, linguistic, and cultural barriers to health care. In 2004, South Cove provided over 100,000 medical encounters to more than 16,000 unduplicated patients. The populations served were predominantly Asian Americans (>96%), and the majority of them were Chinese Americans, based on the patients' self-reported ethnicity. Subjects of this study were required to be able to read and understand the questions either in English or in Chinese. Patients with unstable medical conditions or who were unable to be interviewed in English or one of the four commonly spoken Chinese dialects, Mandarin, Cantonese, Taiwanese and Toisanese, were excluded.

Procedure

This study is part of a Massachusetts General Hospital (MGH) research project funded by the National Institute of Mental Health (MH67085-01A1) to provide culturally sensitive collaborative treatment to depressed Chinese Americans in primary care. The research project has been approved both by the MGH Human Research Committee and the South Cove Human Research Committee, and all subjects participated voluntarily after informed consent was obtained. Since 2001, the primary care clinic in South Cove has implemented depression screening as part of its quality improvement measures. All adult patients who attended the clinic were asked to fill out the Chinese Bilingual (English and Chinese) PHQ-9 (CB-PHQ-9). Support staff at the clinic put down the patient's name and medical record number on the questionnaire for identification, and provided simple instructions regarding how to complete

and return the questionnaire. Patients were informed that they did not need to fill out the CB-PHQ-9 if they had done so within the last three months. Added to the CB-PDQ-9 were two questions asking the patients' permission to be contacted if their answers suggested the possibility of depression and whether they needed urgent psychiatric assistance. The principal investigator (A.Y.), a licensed psychiatrist, was available throughout the study to the primary care staff, to assess patients' suicidal risks when necessary.

Patients who screened positive (defined as CB-PHQ-9 ≥ 15 to identify patients with moderate to moderately severe depression requiring medication treatment), and registered as being Chinese in South Cove's computer system received a phone call from the principal investigator (P.I.) (A.Y.), who informed patients of the results of the screening and recommended that patients receive a clinical assessment. Patients were also informed the opportunity to participate in an NIMH funded study to improve cultural sensitivity treatment for patients with MDD. Patients who were interested in participating in the study and completed an informed consent form were interviewed by the P.I. using the Chinese-Bilingual Structured Clinical Interview for DSM-IV, patient version (CB-SCID-I/P)(12) to confirm the diagnosis of MDD.

Eight percent of the patients who scored between 1 and 14 on the CB-PHQ-9 who indicated their willing to be contacted were randomly selected for a telephone interview using the MDD module of the CB-SCID-I/P. Primary care physicians of patients found to have MDD were notified of the diagnosis in writing, and patients with MDD were encouraged to discuss treatment options with their primary care physicians.

Instruments

I. The Chinese Bilingual Patient Health Questionnaire-9 (CB-PHQ-9): The PHQ-9 is the depression module of the Patient Health Questionnaire (PHQ), a criteria-based instrument. The PHQ-9 consists of one item for each of the nine criterions upon which the diagnosis of DSM-IV major depressive disorders are based. In an earlier PHQ-9 validation study (13), Kroenke et al. used PHQ-9 ≥ 10 as the threshold, and showed that the instrument had a sensitivity of 88% and a specificity of 88% for identifying MDD. Our team translated the PHQ-9 into Chinese, and revised it several times with input from our bilingual clinicians. We added the Chinese translated version of the PHQ-9 to the original questionnaire to compile the Chinese Bilingual PHQ-9 (CB-PHQ-9).

II. The Chinese-Bilingual Structured Clinical Interview for DSM-IV, patient version (CB-SCID-I/P): The Structured Clinical Interview for DSM-IV, patient version was translated into Chinese, and was validated by researchers in China (12). For patients who screened positive (CB-PHQ-9 ≥ 15), CB-SCID-I/P interviews were performed by the principal investigator (A.Y.), who is a native Chinese-speaking psychiatrist with formal SCID training. Questions on different Axis I disorders were asked exactly as translated, and each was based on the individual criteria from the DSM-IV. Confirmation diagnosis for negatively screened patients were established by phone-calls to randomly selected patients with CB-PHQ-9 scores of 14 or less, by the principal investigator or a master-level clinician under the supervision of the principal investigator using the MDD module of the CB-SCID-I/P.

Training and Reliability Assessments—The P.I. has been extensively trained in the use of the SCID-I/P through the use of videotapes and live interviews of patients. A recent assessment of the degree of inter-rater reliability among clinicians of the Depression Clinical and Research Program (DCRP) at the Massachusetts General Hospital in diagnosing MDD with the SCIDI/P has yielded kappa > 0.75. The master-level clinician received training in the use of the SCID-I/P from the P.I., and the rating procedures for all the diagnostic interviews were reviewed during biweekly research meetings to reduce the potential for interviewer drift.

Data Analyses

In this study, the CB-SCID-I/P interview results were used as the standard by which the CB-PHQ-9 was evaluated. Psychometric analyses of the CB-PHQ-9 were performed to review the internal consistency of the scale. Correlation analyses were performed to explore the relationship between the scores of individual items and the total scores of the CB-PHQ-9. Subjects with and without MDD (based on the CB-SCID-I/P) were compared in their had their mean scores of each of the CB-PHQ-9 items, and the total CB-PHQ-9 scores using T-tests. All statistical tests performed were two-sided and the significance level was set at p≤0.05.

To evaluate the validity of the CB-PHQ-9 as a screening instrument, the sensitivity, specificity, positive predictive power and negative predictive value were computed. Sensitivity is the chance that the screening instrument recognizes MDD cases; specificity is the chance that the screening instrument recognizes non-MDD cases; positive predictive value is the chance that people who are screened positive by the screening instrument actually have MDD; and negative predictive value is the chance that people who are screened negative by the screening instrument do not have MDD. Receiver Operating Characteristic (ROC) curve analysis was performed to assess the overall accuracy of the screening instruments. Using various cut-off scores, patients were categorized as being cases and non-cases according to the screening instrument, with a pair of sensitivity and specificity values at each cut-off score. The area under the ROC curve is calculated by plotting sensitivity on the Y axis and "1-specificity" on the X axis. The area under ROC curve of 1.0 indicates a perfect instrument and an area under the ROC curve of 0.5 means that the instrument performs no better than chance alone for case recognition (14).

RESULTS

During the study, 2,620 CB-PHQ-9 were collected, of which 2,035 (78%) were completed, and 585 (22%) were blank/incomplete questionnaires from patients who declined participation, or patients who had filled out a questionnaire in the past three months. Ninety-five patients received more than one screening in a three month-period. For these patients, only the results from the initial screening were included; subsequently, 1,940 unduplicated subjects remained in the analyses. Based on the order of screening, a sub-sample of 92 patients (5% of the 2,035 patients who completed the questionnaire, eight patients were removed since they were screened more than once) were matched to 58 patients (10% of the 585 subjects) who did not complete the questionnaire, and the two groups were compared on their demographics; no statistically significant differences were found in the two groups in their mean ages (50±17 vs 48±17; n.s.), and proportions of being female (79% vs 66%; n.s.). Using CB-PHQ-9 ≥ 15 as the cut-off score, 72 (3.7%) patients screened positive for having moderate to moderately severe MDD; 37 (51%) of them agreed to come to the clinic to be interviewed with the CB-SCID-I/P, of whom 34 (92%) were diagnosed with MDD (figure 1). The positively screened patients who were interviewed and those who declined interview did not differ in their mean ages (51±20 vs. 48±19; n.s.), proportions of being female (70% vs. 53%; n.s.), and mean total CB-PHQ-9 scores (19±3.4 vs 18±3.3; n.s.). Among the 1,868 patients who screened negative, 147 (8%) were randomly selected, based on the date of screening and order of the study number assigned, and were interviewed with the MDD module of the SCID-I/P during a phone-interview; 139 (95%) of whom had no MDD diagnosis.

In analyzing the internal consistency of the CB-PHQ-9 scale for patients with and without MDD, the reliability coefficient (alpha) was computed and found to be 0.91. The correlations between individual items of the CB-PHQ-9 and the total scores ranged from 0.52 to 0.85, and all correlations were significant at the 0.01 level (2-tailed)(Table 1). Patients with MDD had significantly higher average scores in each of the CB-PHQ-9 items and the total CB-PHQ-9 scores, when compared to patients with no MDD (table 2). The distribution of total CB-PHQ-9

scores of non-MDD patients skewed heavily to the left (lower scores), while those of MDD patients skewed heavily to the right (higher scores), with a small degree of overlap in the range from 12-20 (Figure 2). The sensitivity, specificity, positive predictive value, and negative predictive value of CB-PHQ-9 were found to be 81%, 98%, 92% and 95% respectively (table 4), and the area under the ROC curve was estimated to be 0.97 (s.e. 0.01) (Figure 3). These findings suggest that the CB-PHQ-9 is an excellent screening instrument for screening for MDD among Chinese Americans.

DISCUSSION

Based on the results, the CB-PHQ-9 is a valid and useful instrument for screening for MDD among Chinese American immigrants. Subjects in this study were restricted to ethnic Chinese since patients at South Cove were predominantly immigrant Chinese Americans. These findings are consistent with those from our previous study with the CBDI, which suggested that depression screening instruments developed in the west work well in detecting MDD among less acculturated Chinese Americans, despite significant differences in illness beliefs about depression across cultures.

The validation of the CB-PHQ-9 is part of a larger research project to recognize MDD and to provide culturally sensitive collaborative treatment to Chinese immigrants with MDD in a primary care clinic. A relatively high cut-off score of CB-PHQ-9 ≥ 15 was chosen so that most of the patients who screened positive had severe enough depressive symptoms to require psychiatric intervention. With the use of this higher cut-off score, the CB-PHQ-9 maintained a reasonably high sensitivity of 81%, indicating that it was able to recognize most patients with MDD in this study.

The CB-PHQ-9, with demonstrated validity, may serve as a practical instrument for depression screening among Chinese Americans in the primary care setting. In addition to having a brief and valid screening instrument, implementation of routine depression screening requires support from the leadership of the clinic and/or health insurance companies to commit the resources needed. These include the need to develop a record keeping information system for alerting the clinic when patients are due for depression screening, and to provide staff to assist patients filling out the CB-PHQ-9, particularly those who have difficulty reading or understanding the questionnaire. In addition, mental health professionals are needed to establish accurate diagnosis for those who screen positive, to explain depression to the patient in a culturally sensitive manner, and to engage patients in treatment.

One of the limitations of this study is that interviewers who established the psychiatric diagnosis were not blinded to the patients' CB-PHQ-9 scores, which might have influenced the confirmation diagnoses. In addition, subjects who initially screened negative received only the MDD module, and not the complete CB-SCID-I/P. For those patients, the presence of medical illnesses and other psychiatric diagnoses might not have been detected; they could have been misclassified as having MDD while they actually had medical problems. In reviewing the results, only 8 (5%) out of the 147 patients who screened negative were given a MDD diagnosis. The problem of creating false-positives from phone interviews, if it occurred, would have been relatively minor. Another limitation is that subjects enrolled in this study were recent Chinese immigrants with low levels of acculturation. The results from this study may not be generalizable to Chinese Americans who are more acculturated, or to other groups of Asian Americans.

CONCLUSION

The CB-PHQ-9 is a valid screening instrument for recognizing MDD with high sensitivity, and excellent specificity. Its accuracy and brevity make it a useful tool for routine depression screening for immigrant Chinese Americans in primary care settings.

ACKNOWLEDGEMENT

Apendix

華人醫務中心 South Cove Community Health Center
個人健康問卷調查表 Patient Health Questionnaire (PHQ-9)

姓名(Name):		性別 (Gender):		年齡 (Age):	
日期 (Date):		藍卡號碼 (MR#)			

_____ 本人最近已經填寫過問卷/不便填寫 (I have filled out one recently/ not interested)

❖ 填完後請交還給護士 (please return to the clinic nurse)

過去的 **2 個星期**內,你有沒有出現以下的徵狀?
Over the last **2 weeks**, how often have you been bothered by any of the following problems?

請仔細閱讀以下問題,在空格內打✓. **Read each item carefully and ✓ your response.**	沒有 **Not at all**	有幾天 **Several days**	一半以上的天數 **More than half the days**	幾乎每天 **Nearly every day**
做事情興趣減少 Little interest or pleasure in doing things	0	1	2	3
心情絕望無助,憂鬱或悲傷 Feeling down, depressed, or hopeless	0	1	2	3
難以入睡,易醒或睡太多 Trouble falling asleep, staying asleep, or sleeping too much	0	1	2	3
覺得疲勞或沒有力氣 Feeling tired or having little energy	0	1	2	3
胃口不好或吃太多 Poor appetite or overeating	0	1	2	3
覺得自己不好,失敗或讓自己,家人失望 Feeling bad about yourself, feeling that you are a failure, or feeling that you have let yourself or your family down	0	1	2	3
有困難集中精神,如看報紙或電視 Trouble concentrating on things such as reading the newspaper or watching television	0	1	2	3
行動或講話很慢到別人都可以注意到, 或是煩躁到靜不下來 Moving or speaking so slowly that other people could have noticed. Or being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
有自殺的念頭或想傷害自己 Thinking that you would be better off dead or that you want to hurt yourself in some way	0	1	2	3

總分(Total): [ ]   =加總(add columns): [ ] · [ ] + [ ]

❖ 需要緊急的協助·? **Do you need to talk to a counselor urgently?** 需要 (Yes) ____ 不需要 (No) ____

❖如果華人醫務中心認為有需要的話,會進一步跟 聯絡,以了解銀是否有憂鬱的傾向
Do you want to be contacted if your answers suggest possibility of depression?

❖ 可以打電話跟我聯絡 (Yes) ____ ; 電話 (Phone#) _____ ;請不要跟我連絡 (No) ____

REFERENCES

1. U.S. Department of Health and Human Services. A Supplement to Mental Health: A Report of the Surgeon General. Author; Rockville, MD: 2001. Mental health care for Asian Americans and Pacific Islanders; p. 107-126.

2. YEUNG, AS.; KAM, R. Illness beliefs of depressed Asian Americans in primary care. In: Georgiopoulos, AM.; Rosenbaum, JF., editors. Perspectives in cross-cultural psychiatry. Lippincott Williams & Wilkins; 2005. p. 21-36.

3. CHUNG H, TERESI J, GUARNACCIA P, et al. Depressive symptoms and psychiatric distress in low income Asian and Latino primary care patients: prevalence and recognition. Community Mental Health Journal 2003;39(1):33–46. [PubMed: 12650554]

4. U.S. Preventive Services Task Force. Screening for Depression: Recommendations and Rationale. Agency for Healthcare Research and Quality; Rockville, MD: May. 2002 http://www.ahrq.gov/clinic/3rduspstf/depressrr.htm37

5. YEUNG AS, HOWARTH S, CHAN R, et al. Use of the Chinese version of the Beck Depression Inventory for Screening Depression in Primary Care. Journal of Nervous and Mental Disease 2002;190:94–99. [PubMed: 11889362]

6. NEASE KE, MALOUIN JM. Depression Screening: A practical strategy. The Journal of Family Practice 2003;52(2):118–126. [PubMed: 12585989]

7. HUANG FY, CHUNG H, KROENKE K, SPITZER RL. Racial and ethnic differences in the relationship between depression severity and functional status. Psychiatr Serv 2006;57(4):498–503. [PubMed: 16603745]

8. GRYPMA L, HAVERKAMP R, LITTLE S, UNUTZER J. Taking an evidence-based model of depression care from research to practice: making lemonade out of depression. Gen Hosp Psychiatry 2006;28(2):101–7. [PubMed: 16516059]

9. ELL K, UNUTZER J, ARANDA M, SANCHEZ K, LEE PJ. Routine PHQ-9 depression screening in home health care: depression, prevalence, clinical and treatment characteristics and screening implementation. Home Health Care Serv Q 2005;24(4):1–19. [PubMed: 16446263]

10. LOWE B, SCHENKEL I, CARNEY-DOEBBELING C, GOBEL C. Responsiveness of the PHQ-9 to Psychopharmacological Depression Treatment. Psychosomatics 2006;47(1):62–7. [PubMed: 16384809]

11. MARTIN A, RIEF W, KLAIBERG A, BRAEHLER E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. Gen Hosp Psychiatry 2006;28(1):71–7. [PubMed: 16377369]

12. KAM ESI, LEUNG CM, CHUNG D, LIU Z, FONG S. The Chinese-bilingual SCID-I/P Project: Stage 1-Reliability for Mood Disorders and Schizophrenia. Hong Kong J Psychiatry 2003;13(1):7–18.

13. KROENKE K, SPITZER RL, WILLIAMS JB. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med Sep;2001 16(9):606–613. [PubMed: 11556941]

14. HANLEY J, MCNEIL BJ. The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. Radiology 1982;143:29–36. [PubMed: 7063747]
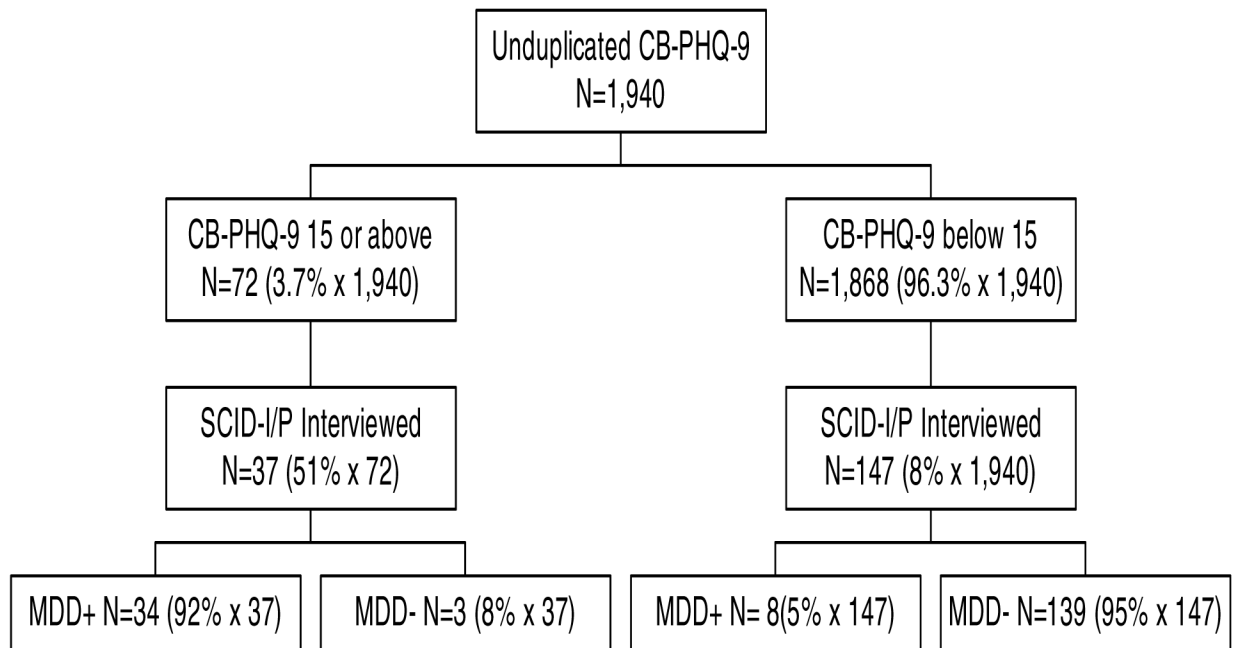
Figure 1.
Screening for Depression using the Chinese Bilingual version of the Patient Health
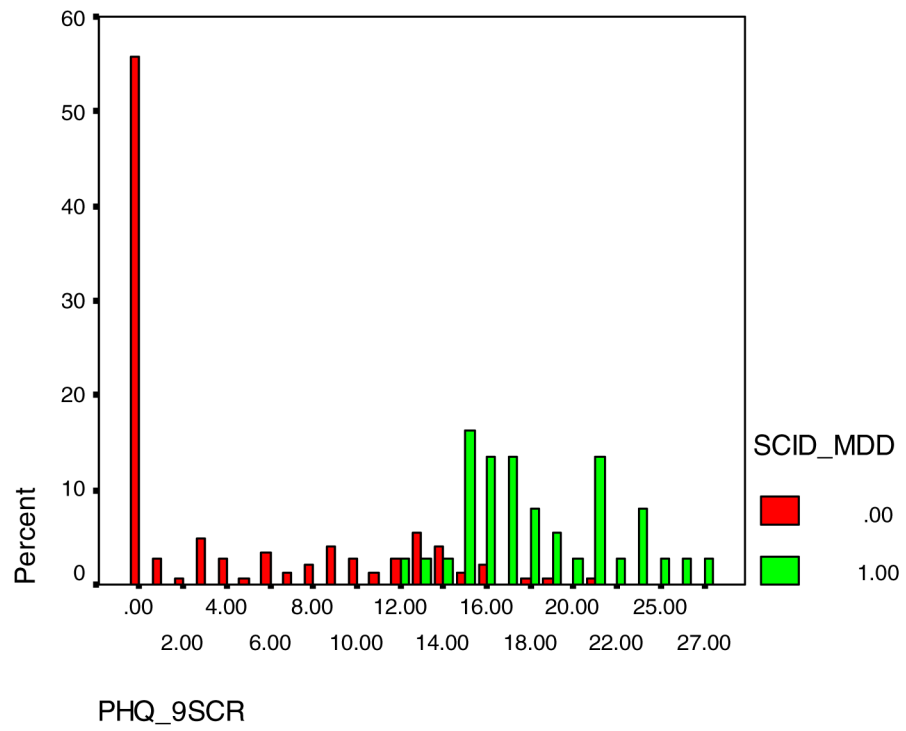Questionnaire (CB-PHQ-9)

Figure 2.
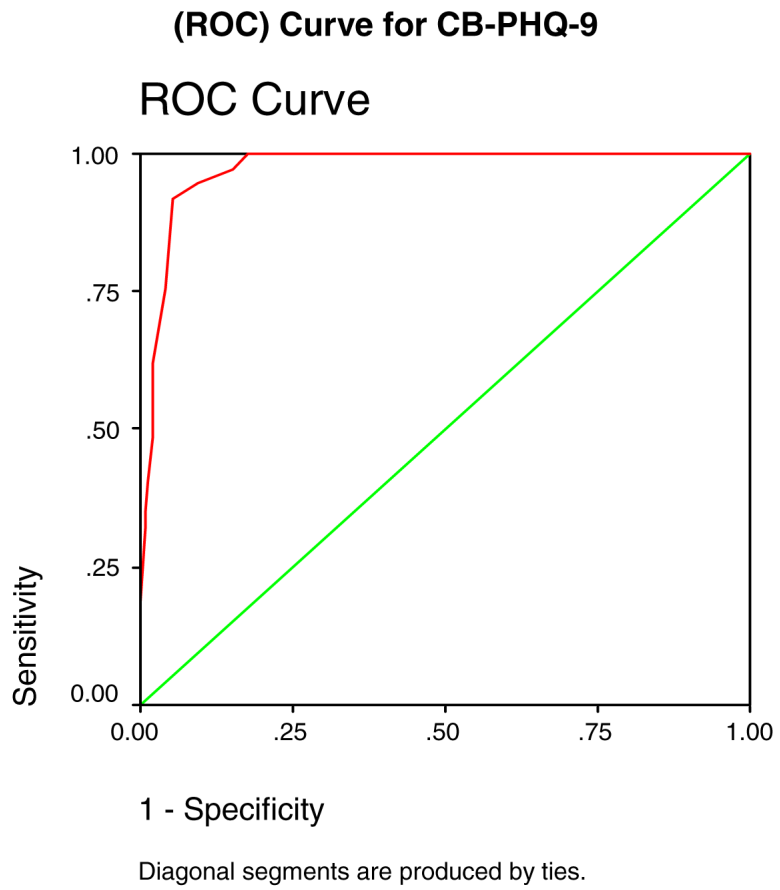Percentage distribution of total CB-PHQ-9 scores of non MDD (n=147) and MDD (n=37) patients

(ROC) Curve for CB-PHQ-9

Figure 3.
Reciever Operating Characteristic (ROC) Curve for CB-PHQ-9

Table 1

Correlation between CB-PHQ-9 item scores and total scores

	Sadness	Insomnia	↓Interest	Guilt	↓ Energy	↓Concentration	↓↑Appetite	↓↑Motor	SI
Total scores	.78**	.85**	.84**	.64**	.84**	.84**	.68**	.81**	.52**

**
Correlation is significant at the 0.01 level (2-tailed)

Table 2

Comparison of symptom scores between MDD and non-MDD patients

Item	non MDD cases (n=147)	MDD cases (n=37)	t-test (2-tail) p value
Sadness	0.44±0.89	2.16±1.12	<0.001
Insomnia	0.80±1.17	2.49±0.96	<0.001
Loss of interest	0.56±0.98	2.49±0.80	<0.001
Guilt	0.21±0.63	1.69±1.33	<0.001
Low energy	0.77±1.13	2.40±1.14	<0.001
Difficulty concentration	0.43±0.93	2.62±0.83	<0.001
Appetite change	0.46±0.96	1.71±1.36	<0.001
Psychomotor retardation/agitation	0.44±0.88	2.08±1.18	<0.001
Suicidality	0.01±0.12	0.86±1.01	<0.001
Total PHQ-9 scores	4.11±5.69	18.38±3.68	<0.001

Table 3

Validity of the CB-PHQ-9 for depression screening among Chinese Americans

		CB-PHQ-9		
		+	-	Total
CB-SCID-I/P	+	34	8	42
	-	3	139	142
Total		37	147	184

Sensitivity = 34/42 = 0.81

Specificity = 139/142 = 0.98

PPV = 34/37 = 0.92

NPV = 139/147 = 0.95

[+]Patients with DSM-IV Major Depressive Disorder

[-]Patients without DSM-IV Major Depressive Disorder