Screening for Major Depression in Asian-Americans:

A Comparison of the Beck and the Chinese Depression Inventory

RUNNING HEAD TITLE: (The Beck and Chinese Depression Inventory)

Albert Yeung[1,2], Nicole Neault[1], Shamsah Sonawalla[1], Shauna Howarth[1],

Maurizio Fava[1], Andrew A. Nierenberg[1]

From the
1Depression Clinical & Research Program, Massachusetts General Hospital

[2]South Cove Community Health Center, Boston

Correspondence: Albert Yeung, M.D., Sc.D.

50 Staniford St., Suite 401

Boston, MA 02114

Office (617) 724-5138

ayeung@partners.org

Screening for Major Depression in Asian-Americans:

A Comparison of the Beck and the Chinese Depression Inventory

Abstract

Objective. This study compares the effectiveness of the Chinese version of the Beck

Depression Inventory (CBDI) and the Chinese Depression Inventory (CDI)

in screening for depression among Chinese Americans.

Method. Five hundred and three Chinese-Americans in primary care were administered

the CBDI and the CDI for depression screening. The results were compared to standard

semi-structured interview.

Results. With empirically determined cutoff scores of the CBDI (≥ 13) and the CDI ($\geq$

16), both instruments have good sensitivity (0.78) and excellent specificity (0.91 and 0.93

respectively). The correlation between the total scores of the two instruments was high

(0.785, $p < 0.01$). The areas under the ROC curve of the CBDI and the CDI were 0.94

and 0.95 respectively and were not significantly different.

Conclusion: When administered by interviewers, the CBDI and the CDI have

comparable effectiveness. Low participation among Chinese-Americans with self-report

measures limits both scales as efficient depression screening instruments. (key words:

Depression, mass screening, primary health care, Asian Americans)

Introduction

Although depression is prevalent among Asian-Americans (1), this population highly under-utilizes the mental health services available to them (2,3). Cultural factors may contribute to low rates of recognition of depression among Asian-Americans, as well as their reluctance to seek help for this illness (4). In European and North American cultures, depression is a psychiatric syndrome characterized by specific affective, cognitive behavioral, and somatic symptoms. In traditional Southeast Asian cultures, there is no equivalent concept of depression as a treatable disease entity (5). Mental illness is highly stigmatized and mental health services are avoided as much as possible by traditional Asian-Americans. When depressed, they usually seek treatment through primary care, where they tend to focus on their physical symptoms and under-report their emotional symptoms (6). The rates of recognition of depression among Asians are low and treatment with antidepressants was uncommon (7). Screening of depression in primary care may be one of the most efficient and effective ways to improve the detection and treatment of depression among less acculturated Asian-Americans.

In cross-cultural psychiatric research, there is a controversy regarding whether to translate existing instruments or to develop culture-specific rating scales for screening psychiatric symptoms. Since depression screening instruments have been validated and extensively studied in the Western countries (8,9), translation of existing instruments appears to be the most sensible method of screening Asian-Americans for depression, without needing to construct and test new instruments for the same purpose.

A significant problem with translating instruments is that symptoms often sound awkward or incomprehensible to indigenous people, even if the wordings are semantically correct. Kinzie (10) maintained that instruments from Western countries, especially self-report inventories, were inappropriate when used in different cultural and ethnic groups. It was not clear whether Asians were willing to report certain affective states and whether they were suppressing the expression of dysphoric feeling. Instead of using the symptoms or distress reported by westerners, Kinzie developed the Vietnamese Depression Scale by using a combination of the Vietnamese perception of lowered mood and behaviors compatible with the Western concept of depression to screen for depression among Vietnamese immigrants.

Similarly, Zheng proposed that in order to measure depression among Chinese, it was necessary to design instruments that incorporate symptoms, words, and idioms specific to their culture (11). Zheng and his colleagues in Hunan Medical University (12) developed the Chinese Depression Inventory (CDI) using words and phrases used by depressed Chinese patients. Zheng et al (12) studied 329 depressed Chinese patients from 24 hospitals across China and compared the psychometric properties of the Chinese version of the Beck Depression Inventory (CBDI) and the CDI. They found that 3 items of the CDI correlated better than the corresponding CBDI items with intensity of depression measured by the total score of the Chinese version of Hamilton Depression Rating. It was concluded that the CDI eliminated cultural biases and was a better instrument than the CBDI to measure severity of depression among Chinese.

While the CDI was shown to have better psychometric properties over the CBDI when applied to Chinese patients, the CBDI has better face validity with items similar to symptoms of major depression described in the DSM-III & IV. Extensive use of the CBDI by in psychiatric research allows comparability of patients' severity of depression across studies. Our earlier study using the CBDI for screening for depression has shown satisfactorily results (13). To date, there are no studies on the effectiveness of the CDI as a screening instrument for recognizing patients with major depression, or whether the CDI is superior to the CBDI for screening Chinese patients with major depression. The present study compares the effectiveness of the CBDI and the CDI for screening for depression among Chinese-Americans in primary care. We hypothesize that the CBDI, a translated instrument with good face validity performs as well as the CDI, an instrument developed with culture-specific items, for screening for depression among Chinese-Americans.

Subjects and Methods

<u>Subjects</u>: Subjects were Asian-Americans who attended South Cove Community Health Center (South Cove), an urban community health center located in the Northeastern part of the U.S. South Cove serves low-income Asian immigrants who face financial, linguistic and cultural barriers to health care. In 1999, South Cove provided 77,811 medical encounters and had 11,751 patients, with 5,897 (50%) from the Adult Medicine (Primary Care) Clinic. The populations served are predominantly Asians (92%); other ethnic groups include African American (1%) and Caucasian (1%). The information for the ethnicity of six percent of patients is not available. The ethnicity of the patients was determined by self-report. Participants of this study had to be able to read Chinese or speak any one the four Chinese dialects including Mandarin, Cantonese, Taiwanese and Toisanese. Participants also had to be 18 years of age or older and they had to sign a written consent to participate in the study. Patients who had unstable medical conditions or were unable to be interviewed were excluded. Patients who were illiterate and had difficulty understanding questions read aloud to them were also excluded from the study.

<u>Study Design and Procedures</u>: The method of convenient sampling was used. Data were collected between May 1998 and November 1999. Patients were asked to fill out the CBDI while they were in the waiting area of the primary care clinic at South Cove. For patients who were illiterate, we offered to read the items of the inventory to them. We discovered very early in the study that most patients were reserved and unwilling to participate. The impersonal approach of leaving patients with the instrument to fill out on their own did not work well with less acculturated Asian-Americans. To make the study

more feasible, we resorted to using a more personal approach. Research assistants sat down next to the patients, introduced himself/herself and explained the nature of the study, and then interviewed the patients using the items of the inventory. Using this modified approach, we were able to engage the patients and we received a much better participation rate. In this study, the CBDI and the CDI were used as research assistant-administered instruments instead of self-report instruments.

All of the patients who scored 16 or higher on the CBDI were scheduled to be interviewed with Structured Clinical Interview for DSM-III-R, patient version (SCID-I/P) to confirm the diagnosis of DSM-III major depression. A portion of the patients who scored below 16 on the CBDI were randomly selected to be interviewed with the depression module of the SCID-I/P. Patients who were interviewed with SCID-I/P or major depression (MDD) module of SCID-I/P were also administered the CDI, for those who agreed.

Instruments

1. Chinese version of the Beck Depression Inventory (CBDI). The BDI is a self-report scale for depression and is widely used to measure the severity of depression for research purposes (8). The CBDI was translated in Chinese and back translated into English by Chinese psychiatrists (11). The procedures of translation and back-translation were continued until the back-translated BDI corresponded closely to the original Beck Depression Inventory. Correlation coefficient of CBDI (using Cronbach's alpha) was found to be 0.85 (12).

2. The Chinese Depression Inventory (CDI). The initial 32 items of the CDI were derived from items of the CBDI, Zung Self-report scales, and the Chinese version of the Hamilton Rating Scale for Depression. Sixteen of the thirty-two items were found to have questionable face validity. These items were replaced and 40 items with equivalent phrases and words commonly used by depressed Chinese patients to express these concepts were gathered during a face validity assessment process (14). The final 48 items of the CDI scale were determined after discussion and revision by 45 Chinese psychiatrists who attended a training course for the application of standard diagnosis and rating scales (15).

3. Structured Clinical Interview for DSM-III-R, patient version (SCID-I/P)(16). The SCID-I/P was translated into Chinese by researchers in the National Cheng Kung University Medical College in Taiwan and was used in a cross-cultural study on Neurasthenia by Zheng et al. (17). SCID-I/P interviews were performed by the principal investigator (ASY), who is a native Chinese-speaking psychiatrist with formal SCID training. He is currently a staff psychiatrist working in the Depression Clinical and Research Program (DRCP) at Massachusetts General Hospital and routinely performs SCID interviews in English. A previous study of the inter-rater reliability with the SCID-I/P among staff psychiatrists at the DCRP yielded a kappa of 0.78 for mood disorders (18).

Statistical Analyses: In this study, the CBDI and the CDI were the screening instruments

for detecting DSM-III-R major depression among Chinese Americans in primary care,

and the SCID-I/P interview results were used as the standard for assessing the two

instruments. The sensitivity, specificity, positive predictive power and negative predictive

value were the indexes used to show the validity of the screening instruments. 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 (19) for case recognition. To

assess whether there is significant difference between the area under the ROC curves for

CBDI and the CDI, the critical ratio z, was calculated (20):

$$Z = \frac{A_1 - A_2}{(SE_1^2 + SE_2^2 - 2r\,SE_1 + SE_2)^{1/2}}$$

The area under the ROC curves and the standard errors associated with the areas were obtained from Wilcoxon statistic. The r is a quantity representing the correlation introduced between the two areas by studying the same sample of patients (19). The degree of the linear relationship between the total scores of the CBDI and the CDI was calculated using the sample correlation coefficient.

Results

Eight hundred and fifteen patients in the primary care clinic waiting area were approached, and 503 (62%) gave consent to participate in the study. All of the 503 patients had CBDI ratings. Among 503 patients, 76 (16%) had a CBDI score $\geq$ 16 (standard cutoff score of CBDI), and 53 of those patients agreed to be interviewed with the SCID. All but five of them were also administered the CDI. Four hundred and twenty-seven patients scored 15 or less on the CBDI. One hundred and twenty-seven of them (30%) were interviewed with the MDD module of the SCID-I/P, from which 69 were also administered the CDI. Of the 117 participants in this study who were administered the CBDI, the CDI and were interviewed with the SCID-I/P, 73 (62.4%) were female and 44 were male (37.6%). The mean age was 50.7 ± 19.0 (ranged from 18 to 84).

Using the standard cutoff scores for both instruments ($\geq$ 16 for the CBDI and $\geq$ 20 for the CDI), the sensitivities, specificities, positive predictive values and negative predictive values of the CBDI and the CDI were 0.67, 0.99, 0.92, 0.94 and 0.67, 0.94, 0.67, 0.94 respectively, with relatively low sensitivities but high specificities. For this study population, using the empirically determined cutoff scores of $\geq$ 13 for the CBDI and $\geq$ 16 for the CDI improved the sensitivity to 0.78 for both screening instruments, while maintaining a specificity > 0.9 (Table 1). The correlation between the total scores of the CBDI and the CDI was high (0.785, p < 0.01). The area under the ROC curve of the CBDI was 0.94 (s.e. 0.028); and the area under the curve for CDI was 0.95 (s.e. 0.0216). With the selection of the optimal thresholds, both scales are excellent

instruments for screening depressed patients. The areas under the ROC curves for CBDI

and CDI were not significantly different (z=0.49, p=0.174)(20)(Figure 1).

Discussion

Our results show that both the CBDI and the CDI, when administered by raters, are effective and comparable for screening for depression among Chinese-Americans in an urban community health center. Both scales had good sensitivity and excellent specificity using empirically determined cutoffs. The total scores of the CBDI and the CDI correlated well with each other. Both scales have excellent discriminatory power reflected by their high areas under the ROC curves. The results support our hypothesis that the CBDI, a translated instrument from Western countries with good face validity for DSM-III-R defined major depression, performs as well as the CDI when administered by native speaking interviewers.

In our earlier study involving many of the same patients, the CBDI had high sensitivity (0.79) and specificity (0.91) for screening for depression using standard cut-off scores (≥16). In this study, both the CBDI and the CDI had only moderate sensitivity (0.67 for both) using standard cutoffs, missing about one third of patients with major depression. The decreased sensitivity of the CBDI (and low sensitivity of the CDI) in this study may be explained by the change in the composition of the participants. It may also reflect the lack of robustness of validity indexes of screening instruments when applied to different populations, as well as the limitations of relying on self-report, the content of which can be affected by the interview setting, emotional states of participants, and the interaction between participants and interviewers. It may be necessary to empirically determine the cut-off score of a screening instrument before applying it to a specific population. The process of finding the best cut-off score to optimize validity

indexes may, by chance, give an overrated performance of the instrument with sensitivity and specificity that are not replicable when the instrument is applied to another population or to a subgroup of the participants.

The participants of this study were mostly recent Asian immigrants with language difficulty and low degree of acculturation. A substantial number of them had low education levels or were illiterate. Soon after we started the study, we discovered the yield was too low to ask them to self-administer the screening instruments. They were markedly reserved and not accustomed to taking part in clinical research. Some of them had difficulty using the Likert scale to report the severity of the symptoms they were experiencing. Instead of handing out the instruments to the patients and leaving them on their own to complete the instruments, research assistants interviewed them with the screening instruments. This approach greatly increased our participation rate.

The use of native speaking interviewers to administer translated instruments may incorporate cultural meanings to items in the instrument which could otherwise sound unnatural or awkward to people from non-western cultures. Previous studies have reported similar difficulties when Asian-immigrants with low education levels were asked to use self-rating scales. The validity of the Vietnamese Depression Scale was better when administered by research assistants compared to when self-rated by patients (21,22).

Use of interviewers to conduct depression screening can be time consuming, costly, and may not be feasible in many primary care settings. In can also introduce biases to the screening process. The interviewers may not use the same wordings each time they talk to a participant. While there is only one written Chinese language, there are many different spoken dialects used by the Chinese populations which may add variations to how the questions are presented. Also, patients may respond differently in the presence of an interviewer; some may present socially desirable answers while others may deny their symptoms that they feel are too embarrassing to report. It would be useful and informative to find out if the rates of depression were different among patients who self-reported their symptoms and patients who were interviewed with the instruments. Since we had adopted the interview approach early in the study, the numbers of self-report patients were too small for meaningful analyses.

Despite high sensitivities, specificities, and the discriminant capability of the CBDI and the CDI in this study, the reluctance of the participants in this study to self-report their symptoms using either instrument is the biggest drawback of using these scales as reliable and efficient instruments for screening for depression among Asian-Americans in primary care. For future depression screening among Asian-Americans in primary care, it may be necessary to develop a briefer screening scale to increase participation rate. Use of interviewers should be reserved as a backup for patients who are illiterate or have difficulty using self-report instruments.

There are other limitations of this study. Not all interviewed participants were administered the full SCID-I/P interview; more than half of the patients were administered only the depression module. With the relatively small number of depressed patients in this study, any misclassification in the SCID-I/P interview could have a substantial effect on the validity indexes of the screening instruments. The participants in this study were recent immigrants with language barriers and low levels of acculturation. The results of this study may not be generalizable to other groups of Chinese-Americans.

The ability to screen and recognize depression is only the first step towards improving mental health services in primary care (23). Many Asian-Americans are unfamiliar with mental illnesses and they tend to avoid mental health services due to their fear and rejection of mental disorders (24). Patients with major depression need to be educated about the need to receive treatment and the benefits that treatment could offer. Mental health services need to be available, easily accessible, and user- friendly. Clinicians treating patients from different cultural backgrounds need to be able to introduce depression in a way both understandable and acceptable within the patients' experiences and their cultural framework. The authors are currently implementing an outpatient psychiatric liaison service in the primary care clinic to provide onsite psychiatric evaluation and treatment, in an attempt to provide easier access for Asian-Americans to mental health services and to decrease cultural barriers to treatment of depressive illnesses.

Acknowledgement

References

1. Takeuchi DT, Chung RCY, Lin KM et al: Lifetime and twelve-month prevalence rates of major depressive episodes and dysthmia among Chinese Americans in Los Angeles. Am J Psychiatry 1998: 155:1407-1414.

2. Bui KV, takeuchi DT: Ethnic minority adolescents and the use of community mental health care services. American Journal of Community Psychology 1992: 20:403-417.

3. Sue S, Morishima J: The Mental Health of Asian-Americans. San Francisco, Jossey-Bass, 1982.

4. Lin, TY, Lin KM. Service delivery issues in Asian-North American communities. Am J Psych 1978: 135: 454-456.

5. Marsella AJ, Sartorius N, Jablensky A, Fenton FR. Cross-cultural studies of depressive disorders: An overview. In Kleinman A, Good B ed. Culture and depression. Berkeley and Los Angeles, California: University of California Press, 1985: 299-324.

6. Kleinman A. Neurasthenia and depression: a study of somatization and culture in China. Cult Med Psych 1982: 6: 117-190.

7. Ustun T.B., Sartorius N. (ed.) Mental Illness in General Health Care: An

International Study. West Sussex, England, John Wiley & Sons, Inc., 1995.

8.   Beck AT. An inventory for measuring depression. Arch Gen Psych 1961: 4:53-61.

9.   Zung WO. A cross-cultural survey of symptoms in dep ression. Am J Psych 1969: 126: 116-121.

10. Kinzie D, Manson SM, Vinh DT, Nguyen TT, Anh B, Pho TN. Development and validation of a Vietnamese-Language Depression Rating Scale. Am J Psychiatry 1982: 139: 1276-1281.

11. Zheng Y, Wei L, Goa L, Zhang G, Wong C. Applicability of the Chinese Beck Depression Inventory. Comprehensive Psychiatry 1988a: 29: 484-489.

12. Zheng YP, Lin KM. Comparison of the Chinese Depression Inventory and the Chinese version of the Beck Depression Inventory. Acta Psychiatr Scand 1991: 84: 531-536.

13. Yeung AS, Howarth S, Chan R, Sonawalla S, Nierenberg A, Fava M. Screening of major depression among Chinese patients in primary care. 153rd Annual Meeting of the American Psychiatric Association. Chicago, Illinois, 2000.

14. Zheng YP, Shu LY. Styles of verbal expression of emotional and physical experiences: a study of depressed patients and normal controls in China. Cult Med Psychiatry 1986: 10: 23-143.

15. Zheng YP. Yang W, Phillips MR, Dai C, Zheng H. Reliability and validity of a Chinese computerized diagnostic instrument. Acta Psychiatr Scand 1988b: 77: 32-37.

16. First MB, Spitzer RL, gibbon M, Williams JBW. Structured Clinical Interview for Axis I DSM-IV Disorders (Version 2.0) - Patient Edition. Biometrics Research Department, New York State Psychiatric Institute, New York, New York 1995.

17. Zheng YP, Lin KM, Takeuchi D, Kurasaki KS, Wong Y, Cheung F. An Epidemiology study of neurasthenia in Chinese-Americans in Los Angeles. Compre Psych 1997: 38: 249-259.

18. Fava M, Alpert JE, Nierenberg AA et al. A validation study of a computerized management system for the diagnosis and treatment of depression. 153rd Annual Meeting of the American Psychiatric Association. Chicago, Illinois, 2000.

19. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982: 143: 29-36.

20. Hanley JA, McNeil BJ. A method of comparing the areas under a receiver operating characteristic (ROC) curve. Radiology 1983: 148: 839-843.

21. Hinton W.L., Nang D, Chen YCJ, Tran CG, Newman TB, Lu F. Screening for major depression in Vietnamese refugees: A validation and comparison of two instruments in a health screening population. J Gen Intern Med 1994: 9: 202-206.

22. Felsman JK, Leong FTL, Johnson MC, Felsman IC. Estimates of psychological distress among Vietnamese refugees: adolescents, unaccompanied minors and young adults. Soc Sci Med 1990: 31 251-6.

23. Karlsson H, Joukamaa M, Lehtinen V. Differences between patients with identified and not identified psychiatric disorders in primary care. Acta Psychiatr Scand 2000: 102: 354-358.

24. Yamamoto J. Therapy for Asian Americans. J Natl Med Assoc 1978: 70: 267-270.

Table 1. Validity of the CBDI and the CDI for screening for depression

		CBDI (≥ 13)			CDI (≥16)	
		case	non-case		case	non-case
SCID	case	14	4		14	4
	non-case	9	90		7	92