ORIGINAL RESEARCH

Psychometric Properties of the Chinese Version of the Clinically Useful Depression Outcome Scale for Patients with Major Depressive Disorder

Hai-Yan Ma[1,2]
Xue-Mei Wang[1,2]
Xiao-Jie Huang[2]
Cheng-Jia Yang[2]
Dong-Fang Sheng[2]
Jing-Jing Yang[2]
Ming-Zhi Xu[2]

[1]The Second School of Clinical Medicine, Southern Medical University, Guangzhou, Guangdong, People's Republic of China; [2]Guangdong Mental Health Center, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, People's Republic of China

Objective: This study aimed to evaluate the psychometric properties of the Chinese version of the Clinically Useful Depression Outcome Scale (CUDOS).

Methods: One hundred ninety patients with major depressive disorder (MDD) according to Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria were recruited to the study. The English version of the CUDOS was translated into Chinese using a forward and backward translation method, which was according to the guidelines of adaptation and validation of instruments in cross-cultural health care research. The Chinese version of the CUDOS, the 17-item Hamilton Rating Scale for Depression (HRSD) and the improved Clinical Global Impression-Severity Scale (iCGI-S) were used to evaluate depressive symptoms in one hundred ninety patients with MDD. One week after the first evaluation, sixteen patients were selected randomly for a second assessment. Reliability and validity tests and receiver operating characteristic curves were performed.

Results: The internal consistency of the CUDOS was 0.95, and the split-half reliability coefficient of the CUDOS was 0.92. The correlation coefficient of the retest in sixteen patients was 0.77 ($P < 0.01$). There was a significant difference in the total score of the Chinese version of the CUDOS between the different levels of depression severity groups ($P < 0.01$). The ability of the CUDOS to identify patients in remission was high (area under ROC curve= 0.97). A cut-off score of 14/15 yielded 90.20% sensitivity and 93.60% specificity when iCGI-S=1.

Conclusion: The Chinese version of the CUDOS is valuable as a brief and reliable instrument to assess depressive symptoms and clinical outcome. The findings suggest that the optimal cut-off score to identify patients in remission was 14/15.

Keywords: MDD, the Chinese version of CUDOS, validity, reliability, cut-off score

Introduction

Major depressive disorder (MDD) is a disabling mental disorder because it leads to impaired social function and decreased quality of life, as well as increased morbidity and mortality.[1,2] In addition to mood symptoms, individuals with MDD experience impairments in physical, occupational and social functioning.[3] High prevalence and recurrence rates of MDD may lead to an increased financial burden on global healthcare system.[4–6] The World Health Organization (WHO) estimates that MDD will be the leading cause of the global burden of disease by 2030.[7] The importance of treating patients until achieving a full remission was emphasized in

Correspondence: Ming-Zhi Xu
Guangdong Mental Health Center, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, No. 123 Huifu West Road, Guangzhou, 510120, Guangdong, People's Republic of China
Tel +86 81888553
Fax +86 81862664
Email mingzhixu@126.com

many studies.[8,9] The residual symptoms in treatment associated with a much greater likelihood of relapse or recurrence of depression.[10]

The outcome of MDD is mainly based on the observations of clinicians and the self-report of patients. There are different tools to measure the levels of depression in the clinical practice. The 17-item Hamilton Rating Scale for Depression (HRSD),[11] and the Montgomery–Asberg depression rating scale (MADRS)[12] are commonly used to evaluate depressive symptoms and the effect of treatment in the field of depression. The current goal of antidepressant therapy is achievement of remission, which is defined solely on the basis of symptom severity scores as follows: ≤7 on the $HRDS_{17}$, ≤10 on the MADRS.[13] However, the work of evaluating the patients' symptoms with these instruments usually needs to be completed by trained psychiatrists and costs a large amount of time, subjective bias among different psychiatrists can affect the outcomes of these assessments.[14] There are also many convenient self-report scales, such as the Patient Health Questionnaire-9 (PHQ-9),[15] the Beck Depression Inventory (BDI),[16] the hospital anxiety and depression scale (HADS).[17] These self-management questionnaires are the most commonly used self-rating scales for measuring depression worldwide. But the scales mentioned above are designed to evaluate patients from the perspective of clinical symptoms, they cannot make global assessments of depressed. The Clinical Global Impression scale (CGI) is a classic instrument for making global assessments. The overall impression during an interview is a main element of psychiatric evaluation, but it is vague and difficult to operationalize.[18] Therefore, The improved Clinical Global Impression Scale (iCGI) was designed to improve the validity of CGI severity and improvement scales in the field of depression.[18] However, none of them is capable of assessing both symptoms and clinical outcomes at the same time, unlike the Clinically Useful Depression Outcome Scale (CUDOS).

The CUDOS was designed to be a brief, quickly scored instrument, within the MIDAS project (Rhode Island Methods to Improve Diagnostic Assessment and Service) to develop useful instruments for clinical practice.[19] The CUDOS is a self-report scale, containing 18 items: 16 items concern the diagnostic criteria of MDD according to the DSM-IV, and the other two items concern the psychosocial damage caused by MDD and the quality of life of the patients.[19] Every item in CUDOS uses 0 to 4 to represent the frequency of these symptoms in the past week (0=not at all true (0 days), 1=rarely true (1–2 days), 2=sometimes true (3–4 days), 3=often true (5–6 days), 4=almost always true (every day), with a total score of 0 to 72. The higher the score is, the more severe the depression is. The CUDOS can provide more complete information on the symptoms of MDD and improve work efficiency. It can also be subjected to rigorous scientific scrutiny and used by other researchers for scientific research.[19]

The CUDOS has been translated into various languages and has satisfactory reliability and validity in different regions and populations, such as Korea and Spain.[14,20] It has been reported that CUDOS has been successfully used to evaluate the depressive symptoms of patients with type 2 diabetes in Taiwan area.[21] However, there is no such previous research on the reliability and validity of the Chinese version of the CUDOS in patients with MDD in China mainland. The purpose of this study was to assess the reliability and validity of the Chinese version of the CUDOS for the evaluation of depression in Chinese patients with MDD.

Methods
Design and Participants
This study was supported by the Guangdong Science and Technology Project (project NO: 2017A020215095) and was approved by the Clinical Research Ethics Committee of Guangdong Provincial People's Hospital (No. GDREC2018156H (R1)). This study was conducted in accordance with the Declaration of Helsinki. Authorization for the use of CUDOS was approved by the original author Mark. Zimmerman. The patients with MDD were recruited from Guangdong Provincial People's Hospital from October 2018 to October 2019. A total of 190 patients were recruited and met the following conditions: (1) all patients were provided written informed consent; (2) they were 18–60 years old; (3) they could read and understand all of the questions in the Chinese version of the CUDOS; and (4) they met the diagnostic criteria of MDD in the DSM-5 during last three months. The exclusion criteria were as follows: (1) suffering from bipolar disorder and other mental disorders; (2) having a history of alcohol or drug abuse in the past year; and (3) suffering from serious physical (medical) conditions. All patients were assessed while receiving their treatment as usual, which could be long-term ongoing or recently prescribed.

Translation of the CUDOS

The translation work started after obtaining the consent of the original author. The CUDOS was translated according to the guidelines of adaptation and validation of instruments in cross-cultural health care research.[22] Firstly, the English version of the CUDOS was translated through the forward translation process by two bilingual and bicultural translators, all of whom were fluent in both Chinese and English. One of them is familiar with colloquial phrases, health care slang and jargon, idiomatic expressions, and emotional terms in common use in the Chinese. Secondly, the two translated versions of the instrument were compared, any ambiguities and questions were discussed and resolved by a discussion group. The final version was reviewed by a professional translator and a scholar of Chinese language studies and was agreed upon by the discussion group. Translation and back-translation of the CUDOS were repeated after state-of-the-art procedures in cross-cultural assessment[22,23] until the discussion group thought that the adapted Chinese version was suitable for Chinese patients with MDD.

Instruments

Clinically Useful Depression Outcome Scale (CUDOS)

The CUDOS is a self-report scale used to evaluate depressive symptoms and the functional aspects of depression. Depressive symptoms, the psychosocial damage caused by MDD and the quality of life were assessed by the Chinese version of the CUDOS.

Hamilton Rating Scale for Depression (HRSD)

The 17-item HRSD has been the most frequently used clinician-rated severity of depression scale in antidepressant efficacy trials.[24] The HRSD has demonstrated satisfactory internal consistency reliability and validity.[25]

improved Clinical Global Impression Scale (iCGI)

The Clinical Global Impression-Severity Scale (CGI-S) is widely used in clinical practice, and the severity of patients' depressive symptoms is evaluated by an 7-point Likert scale (from 1 = Normal to 7= Among the most extremely ill patients).[26] The iCGI is designed to address a broad spectrum of depressive disorders, and improves the response format for the Clinical Global Impression severity scale in depression.[18]

Procedure

All of the participants underwent a clinical assessment by a trained psychiatrist using the DSM-5 criteria and those meeting the diagnostic criteria for MDD according to the DSM-5 were referred to this study. After entering the study group, their general demographic data (the patient's age, gender, occupation, education level, family history of mental illness, age of first onset, marital status, frequency of onset, etc.) were obtained. Then, the Chinese versions of the HRSD, iCGI and CUDOS were completed. The participants complete the instruments in a quiet room. Sixteen participants were randomly chosen to return to the hospital seven days after the initial test. We conducted an analysis with the entire sample, which included patients who met the full criteria for MDD as well as patients who were in partial and full remission.

Data Analyses

All analyses were performed using IBM SPSS version 25.0. The continuous variables are described by the mean ± standard deviation (SD), and the categorical variables were described by frequency and percentage. Internal consistency reliability was evaluated using Cronbach's alpha. For clinical applications, much higher values of Cronbach's alpha are needed.[27] Test-retest reliability was estimated with the intraclass correlation coefficient (ICC). The CUDOS dimensionality was assessed with principal component analysis (PCA). The Scree test, parallel analysis and interpretability of the simple structure were used to determine the number of components to be retained. A P-value less than 0.05 was considered to indicate a statistically significant difference between the data sets.

Reliability Testing

Reliability refers to the stability and consistency of the results measured by the scale. In this study, internal consistency reliability (Cronbach's alpha, split-half reliability) and test-retest reliability coefficient of the total scores of the CUDOS were selected to test the stability of the scale.

Internal Consistency

The Cronbach's alpha coefficient was used to evaluate the internal consistency of the total scores scale of the CUDOS. Cronbach's alpha coefficient between 0.90 and 0.95 indicates that the internal consistency of the scale is

desirable, between 0.80–0.90 indicates that the consistency of the scale is good, and between 0.70–0.80 indicates that the consistency of the scale is regarded as satisfactory.[27]

Split-and-Half Reliability Coefficient

Split-and-Half Reliability refers to the half-and-half correlation coefficient between the first half part scores and the remaining part scores of the Clinically Useful Depression Outcome Scale.

Test-Retest Reliability Coefficient

Test-retest reliability is used to test the cross-time stability and consistency of the scale. Test-retest reliability of the CUDOS was evaluated with the intraclass correlation coefficient (ICC). ICC was interpreted according to the ranges of clinical significance proposed by Cicchetti (1994).[28] Test-retest reliability was examined in a sample of 16 participants. They repeating the test after an interval of 7 days.

Validity Testing

The construct validity was investigated using principal component analysis (PCA). The Kaiser–Meyer–Olkin measure and Bartlett's spherical test was used to evaluate the data are suitable for PCA. In this study, Spearman correlation analysis was conducted to determine the correlation between the Chinese version of the CUDOS scores and the scores for the Chinese version of the HRSD. The

ability of the CUDOS to discriminate between different levels of depression severity was investigated based on the rating of iCGI-S. An analysis of variance and post-hoc comparison of Tukey's honestly significant difference (HSD) test was used to evaluate the known-groups construct validity.

Sensitivity and Specificity Testing

Sensitivity and specificity were evaluated by assessing the receiver operating characteristic (ROC) curve to obtain the optimal cutoff score when screening for major depression.[29] The Youden Index for any potential cut-off score is defined as the sum of the sensitivity and specificity (expressed as probabilities) of the scale at that point minus 1. The cut-off is selected as the point with the highest Youden Index.[30,31]

Results

Characteristics of the Participants

A total of 203 patients with MDD were approached for this study and completed their interviews. After the analysis of the selection criteria, 13 patients had to be excluded either because they had not been diagnosed with MDD during the last three months (N=10) or because they could not complete the questionnaire successfully (N=3). Thus, the final valid sample for this study was made up of 190 patients. Their

Table 1 Demographical Characteristics and Clinical Data of the Patients

Variables	Total Sample (n=190)	Depression (iCGI-S >1, n=143)	Remission (iCGI-S=1,n=47)	P-value
Age(years)	29.85±10.36	28.36±9.19	34.38±12.34	P=0.003[a]
Age of onset(years)	26.44±9.11	25.27±9.02	29.96±11.64	P=0.005[a]
Gender, n (%) Male Female	58(30.5) 132(69.5)	39(27.3) 104(72.7)	19(40.4) 28(59.6)	χ^2=136.63 df=2 P=0.000[b]
Marital status, n (%) Married Single	71(37.4) 119(62.6)	46(32.2) 57(67.8)	25(53.2) 22(46.8)	χ^2=0.94 df=1 P=0.332[b]
Education(years)	13.61±3.48	13.79±3.27	13.06±4.06	P=0.216[a]
HRSD scores	15.60±10.09	19.67±8.12	3.21±2.57	P=0.000[a]
CUDOS scores	28.01±18.47	35.45±14.80	5.34±5.29	P=0.000[a]

Notes: [a]Independent Samples *t*-test; [b]Chi-square Test; Continuous variables are expressed as mean ± SD and percentages for dichotomous ones referred to single groups are reported into parentheses.
Abbreviation: SD, standard Deviation.

Table 2 Internal Consistency and Corrected Item-Total Correlation of CUDOS Items

CUDOS Items	Cronbach's Alpha if Item Deleted	Corrected Item-Total Correlation
1. Depressed mood	0.945	0.883
2. Decreased interest in usual activities	0.946	0.845
3. Decreased appetite	0.949	0.639
4. Increased appetite	0.955	0.098
5. Insomnia	0.952	0.527
6. Hypersomnia	0.955	0.302
7. Psychomotor agitation	0.948	0.725
8. Psychomotor retardation	0.948	0.705
9. Decreased energy	0.946	0.827
10. Guilt	0.947	0.771
11. Worthlessness	0.946	0.808
12. Impaired concentration	0.947	0.795
13. Indecisiveness	0.947	0.790
14. Death wishes	0.948	0.694
15. Suicidal thoughts	0.949	0.663
16. Hopelessness	0.946	0.814

sociodemographic and clinical descriptive variables are shown in Table 1. The mean score on the 17-item Hamilton depression scale for the 190 depressed patients was 15.60 (SD=10.09). The mean score of the CUDOS was 28.01 (SD=18.47).

Reliability Analysis

The Cronbach's alpha coefficient of the total score of the CUDOS was 0.95 and the split-half reliability coefficient of the CUDOS was 0.92, respectively, suggesting excellent internal consistency reliability of the Chinese version of the CUDOS. Removal of any of the CUDOS items did not increase the Cronbach's alpha coefficient (Table 2). The intraclass correlation coefficient (ICC) for the total score

of the CUDOS was 0.77 (95% CI 0.449–0.911, $P < 0.01$), which shows excellent test-retest reliability.

Validity Analysis

The results of the Kaiser–Meyer–Olkin measure of sampling adequacy were excellent (0.94), and Bartlett's spherical test showed significant (χ^2=2962.81, P<0.01), which indicates that the data are suitable for PCA. The ratio between the eigenvalues of the first and second components was 7.27 (10.26/1.41), which exceeds the critical value of 4, a criterion usually used as evidence of unidimensionality.[32] This one-factor structure explained 56.99% of the variance. All factor loading values were positive, statistically significant (P<0.05), and ranged from being good to excellent.

The Pearson correlation coefficient between the HRSD and the CUDOS was 0.90 ($P < 0.01$). Table 3 displays the CUDOS mean scores' distribution according to the depression severity levels as measured by the iCGI-S. The mean CUDOS scores of the different depression severity levels were 7.73±7.61, 27.29±9.28, 37.71±10.08, 45.64±7.21, 52.25±8.06, respectively. The five-groups analysis of variance was significant (F=173.74; df=4; P<0.01) and the differences among these groups were significant using Tukey's HSD test.

Sensitivity and Specificity Analysis

A receiver operating curve is a plot of a measure's sensitivity versus one minus the specificity at each cutoff score.[33] When iCGI-S=1 was the "gold standard" definition of remission, the area under the ROC curve of CUDOS was 0.97 (95% CI: 0.950–0.990, $P < 0.01$, Figure 1). A score of 14/15 as the optimal cutoff point was suggested when screening for major depression. At this cutoff point, the sensitivity was 90.20% and the specificity was 93.60% in our sample (Table 4).

When we use the HRSD score less than or equal to seven as used as the criterion for non-depression,[34] the area under the ROC curve of CUDOS was 0.97 (95% CI: 0.953–0.993, $P < 0.01$, Figure 1). This high AUC shows that CUDOS can

Table 3 The CUDOS Mean Scores' Distribution According to Depression Severity Levels as Measured by the iCGI-S

	iCGI-S=1–2 (n=66)	iCGI-S=3 (n=34)	iCGI-S=4 (n=45)	iCGI-S=5 (n=25)	iCGI-S=6–7 (n=20)	Tukey's HSD Test
CUDOS	7.73±7.61	27.29±9.28	37.71±10.08	45.64±7.21	52.25±8.06	(1–2)<(3),(1–2)<(4), (1–2)<(5),(1–2)<(6–7), (3)<(4), (3)<(5), (3)<(6–7),(4)<(5),(4)<(6–7)

Note: Continuous variables are expressed as mean ± SD.
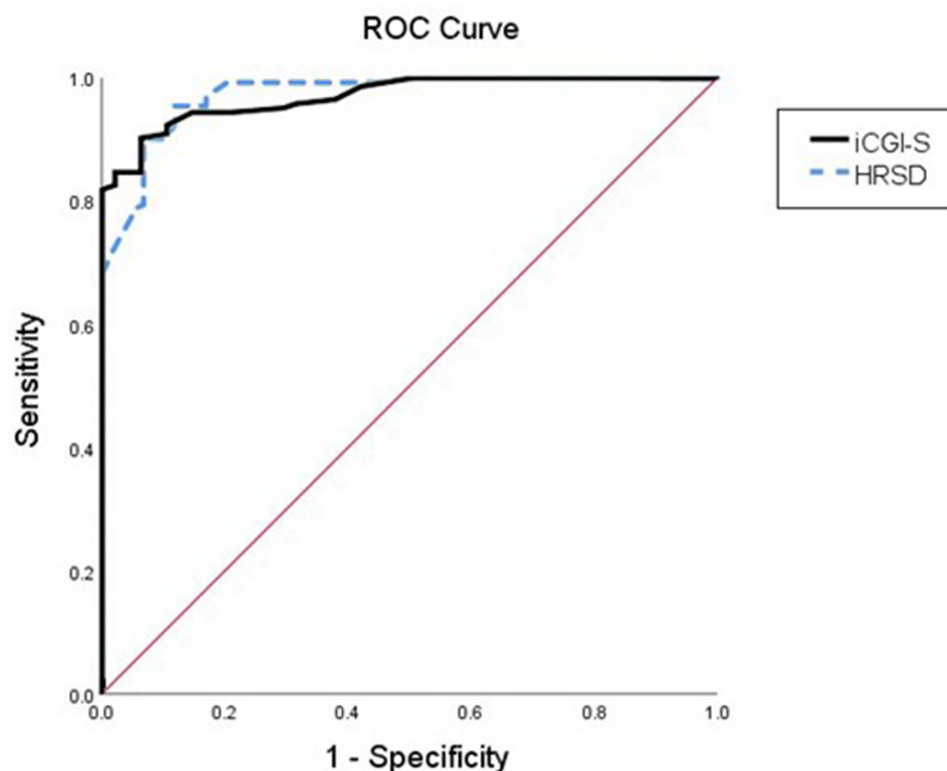**Abbreviation:** SD, standard deviation.

Figure 1 The receiver operating characteristic (ROC) curve of the Chinese version of CUDOS for identifying patients in remission.
Notes: When iCGI=1 was the "gold standard" definition of remission, the area under the ROC curve of CUDOS was 0.97 (95% CI: 0.950–0.990, *P* < 0.01); When HRSD≤7 was used as the criterion for remission, the area under the ROC curve of CUDOS was 0.97 (95% CI: 0.953–0.993, *P* < 0.01).
Abbreviations: HDRS, The 17-item Hamilton Rating Scale for Depression; iCGI-S, The improved Clinical Global Impression-Severity Scale.

quite accurately distinguish between depressed and nondepressed patients. The optimal cutoff score to discriminate between depressed and nondepressed patients was 14/15, with a sensitivity of 95.40% and specificity of 88.10%.

Discussion

The results of this study showed that the Chinese version of the CUDOS is reliable, valid instrument identifying depression patients in remission in a Chinese clinical setting. It retains the same psychometric properties as the original version. The internal consistency of the CUDOS was 0.95 and the split-half reliability coefficient of the CUDOS was 0.92, which indicates the agreement among the individual components of the scale was very good. The test-retest reliability of the scale was 0.77, demonstrating an excellent temporal stability. Regarding to its internal structure, the factor analysis of the CUDOS identified a single-factor structure, which was consistent with the results of previous studies.[19,20] This findings implied that the single-factor model of this scale was stable in different culture populations.[35] The Pearson correlation coefficient

between the HRSD and the CUDOS was 0.90, which showed that the CUDOS was more highly correlated with the HRSD than with measures of the other symptom domains. The known-groups construct validity of the Chinese version of the CUDOS seems to be excellent in discriminating between different depression severity levels patients. The area under the receiver operating characteristics curve of 0.97 indicates that the CUDOS has excellent properties for use as a screening instrument in the identification of MDD.

The cut-off score of the CUDOS for Chinese patients in this study was lower than the previously recommended value.[33] Another study reported that a score of 19/20 was the optimal cutoff score for identifying MDD, which is consistent with that of the original CUDOS cut-off score for identifying mild depression and remission.[36] The CUDOS has demonstrated satisfying validity and reliability to support its use for the assessment of the severity of depression symptoms in patients diagnosed with type 2 diabetes mellitus in routine clinical practice.[21] Our study evaluated the Chinese version of the CUDOS is useful for

Table 4 Sensitivity, Specificity, and Youden Index at Various Cut-off Points of the CUDOS

CUDOS Cutoff Score	Sensitivity	1-Specificity	Specificity	Youden Index
0.500	1.000	0.936	0.064	0.064
1.500	1.000	0.745	0.255	0.255
2.500	1.000	0.553	0.447	0.447
3.500	1.000	0.511	0.489	0.489
4.500	0.986	0.426	0.574	0.560
5.500	0.965	0.383	0.617	0.582
6.500	0.958	0.319	0.681	0.639
7.500	0.951	0.298	0.702	0.653
8.500	0.944	0.213	0.787	0.731
9.500	0.944	0.149	0.851	0.795
11.000	0.923	0.106	0.894	0.817
13.000	0.909	0.106	0.894	0.803
14.500	0.902	0.064	0.936	0.838
15.5000	0.888	0.064	0.936	0.824
16.5000	0.874	0.064	0.936	0.810
18.0000	0.846	0.064	0.936	0.782
19.5000	0.846	0.021	0.979	0.825
21.0000	0.825	0.021	0.979	0.804
22.5000	0.811	0.000	1.000	0.811
23.5000	0.797	0.000	1.000	0.797
24.5000	0.769	0.000	1.000	0.769
25.5000	0.755	0.000	1.000	0.755
26.5000	0.741	0.000	1.000	0.741
27.5000	0.734	0.000	1.000	0.734
28.5000	0.699	0.000	1.000	0.699
29.5000	0.664	0.000	1.000	0.664
30.5000	0.622	0.000	1.000	0.622
31.5000	0.594	0.000	1.000	0.594
32.5000	0.580	0.000	1.000	0.580
33.5000	0.573	0.000	1.000	0.573
34.5000	0.538	0.000	1.000	0.538
35.5000	0.517	0.000	1.000	0.517
36.5000	0.497	0.000	1.000	0.497
37.5000	0.469	0.000	1.000	0.469
38.5000	0.441	0.000	1.000	0.441
39.5000	0.420	0.000	1.000	0.420
40.5000	0.413	0.000	1.000	0.413
41.5000	0.399	0.000	1.000	0.399
42.5000	0.350	0.000	1.000	0.350
43.5000	0.315	0.000	1.000	0.315
44.5000	0.301	0.000	1.000	0.301
45.5000	0.273	0.000	1.000	0.273
46.5000	0.238	0.000	1.000	0.238
47.5000	0.231	0.000	1.000	0.231
48.5000	0.203	0.000	1.000	0.203
49.5000	0.182	0.000	1.000	0.182
50.5000	0.175	0.000	1.000	0.175
51.5000	0.161	0.000	1.000	0.161

(*Continued*)

Table 4 (Continued).

CUDOS Cutoff Score	Sensitivity	1-Specificity	Specificity	Youden Index
52.5000	0.154	0.000	1.000	0.154
53.5000	0.126	0.000	1.000	0.126
54.5000	0.112	0.000	1.000	0.112
55.5000	0.077	0.000	1.000	0.077
56.5000	0.063	0.000	1.000	0.063
57.5000	0.056	0.000	1.000	0.056
58.5000	0.042	0.000	1.000	0.042
59.5000	0.035	0.000	1.000	0.035
60.5000	0.021	0.000	1.000	0.021
61.5000	0.014	0.000	1.000	0.014
64.0000	0.007	0.000	1.000	0.007
67.0000	0.000	0.000	1.000	0.000

measuring the symptoms and outcomes of MDD. Our findings suggested that the same cut-off score might not be appropriate for all settings. There are different cut-off values under different regions.[21] Based on our findings, we recommend using a cut-off score of 14/15 for CUDOS when identifying patients in remission among MDD patients in China.

The HRSD has been one of the most widely used outcome measures in the clinical setting and antidepressant efficacy trials (AETs).[37] Self-applied scales such as the Beck Depression Inventory (BDI), and PHQ-9[7,38] have been successfully used in evaluating psychiatric outpatients as well as in clinical trials. Compared with these depression scales, CUDOS has advantages in evaluating not only psychosocial impairment and quality of life but also clinical symptoms of MDD. More than anything, patients reported that it took less time and reduced the burden on the evaluators.

Several limitations of the study should be noted. Firstly, the main limitation of this study lies in the potential sampling bias due to single-center cohort recruitment. Thus, the generalizability of CUDOS to patients with different socio-demographic or clinical characteristics needs to be determined in future works. Secondly, bipolar depression and other mental disorders were excluded in this study. Therefore, our findings are only applicable to assess the clinical outcome of patients with MDD. Further research should be conducted in different disease patterns to expand the application and confirm the validity. Thirdly, this study provides only cross-sectional data. It is not clear whether CUDOS remains sensitive over time to reflect the outcome of MDD.

Conclusion

In conclusion, the results of this study suggested that the Chinese version of the CUDOS is an effective and reliable self-administered questionnaire to measure depression outcomes in the Chinese population. The optimal cut-off scores of 14/15 are recommended for identifying patients in remission. The Chinese version of the CUDOS appears to be a useful measurement for evaluating the remission and the outcome of MDD in clinical practice.

Acknowledgments

Disclosure

The authors declare no potential conflicts of interest with respect to the authorship or publication of this paper.

References

1. Culpepper L. Understanding the burden of depression. *J Clin Psychiatry*. 2011;72(6):e19. doi:10.4088/JCP.10126tx1c

2. Vandeleur CL, Fassassi S, Castelao E, et al. Prevalence and correlates of DSM-5 major depressive and related disorders in the community. *Psychiatry Res*. 2017;250:50–58. doi:10.1016/j.psychres.2017.01.060

3. Oluboka OJ, Katzman MA, Habert J, et al. Functional recovery in major depressive disorder: providing early optimal treatment for the individual patient. *Int J Neuropsychopharmacol*. 2018;21(2):128–144. doi:10.1093/ijnp/pyx081

4. Charlson FJ, Stapelberg NJ, Baxter AJ, Whiteford HA. Should global burden of disease estimates include depression as a risk factor for coronary heart disease? *BMC Med*. 2011;9:47. doi:10.1186/1741-7015-9-47

5. Ten Doesschate MC, Bockting CL, Koeter MW, Schene AH. Prediction of recurrence in recurrent depression: a 5.5-year prospective study. *J Clin Psychiatry*. 2010;71(8):984–991. doi:10.4088/JCP.08m04858blu

6. Vuorilehto MS, Melartin TK, Isometsa ET. Course and outcome of depressive disorders in primary care: a prospective 18-month study. *Psychol Med*. 2009;39(10):1697–1707. doi:10.1017/S0033291709005182

7. Zhang YL, Liang W, Chen ZM, et al. 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(4):268–275. doi:10.1111/appy.12103

8. Keller MB. Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. *JAMA*. 2003;289(23):3152–3160. doi:10.1001/jama.289.23.3152

9. Romera I, Perez V, Menchon JM, Delgado-Cohen H, Polavieja P, Gilaberte I. Social and occupational functioning impairment in patients in partial versus complete remission of a major depressive disorder episode. A six-month prospective epidemiological study. *Eur Psychiatry*. 2010;25(1):58–65. doi:10.1016/j.eurpsy.2009.02.007

10. Zimmerman M, Posternak MA, McGlinchey J, Friedman M, Attiullah N, Boerescu D. Validity of a self-report depression symptom scale for identifying remission in depressed outpatients. *Compr Psychiatry*. 2006;47(3):185–188. doi:10.1016/j.comppsych.2005.07.004

11. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56–62. doi:10.1136/jnnp.23.1.56

12. Zimmerman M, Posternak MA, Chelminski I. Derivation of a definition of remission on the Montgomery-Asberg depression rating scale corresponding to the definition of remission on the Hamilton rating scale for depression. *J Psychiatr Res*. 2004;38(6):577–582. doi:10.1016/j.jpsychires.2004.03.007

13. Sheehan DV, Nakagome K, Asami Y, Pappadopulos EA, Boucher M. Restoring function in major depressive disorder: a systematic review. *J Affect Disord*. 2017;215:299–313. doi:10.1016/j.jad.2017.02.029

14. Jeon SW, Han C, Ko YH, et al. Measurement-based treatment of residual symptoms using clinically useful depression outcome scale: Korean validation study. *Clin Psychopharmacol Neurosci*. 2017;15(1):28–34. doi:10.9758/cpn.2017.15.1.28

15. Kroenke K, Spitzer RL, Williams JBW. DSW2 The PHQ-9. *J Gen Intern Med*. 2001;16(9):606–613. doi:10.1046/j.1525-1497.2001.016009606.x

16. Hautzinger M. [The Beck Depression Inventory in clinical practice]. *Nervenarzt*. 1991;62(11):689–696.

17. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983;67(6):361–370. doi:10.1111/j.1600-0447.1983.tb09716.x

18. Kadouri A, Corruble E, Falissard B. The improved Clinical Global Impression Scale (iCGI): development and validation in depression. *BMC Psychiatry*. 2007;7:7. doi:10.1186/1471-244X-7-7

19. Zimmerman M, Chelminski I, McGlinchey JB, Posternak MA. A clinically useful depression outcome scale. *Compr Psychiatry*. 2008;49(2):131–140. doi:10.1016/j.comppsych.2007.10.006

20. Trujols J, Feliu-Soler A, de Diego-adelino J, et al. A psychometric analysis of the clinically useful depression outcome scale (CUDOS) in Spanish patients. *J Affect Disord*. 2013;151(3):920–923. doi:10.1016/j.jad.2013.08.007

21. Hsu LF, Kao CC, Wang MY, Chang CJ, Tsai PS. Psychometric testing of a Mandarin Chinese Version of the clinically useful depression outcome scale for patients diagnosed with type 2 diabetes mellitus. *Int J Nurs Stud*. 2014;51(12):1595–1604. doi:10.1016/j.ijnurstu.2014.05.004

22. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract*. 2011;17(2):268–274. doi:10.1111/j.1365-2753.2010.01434.x

23. Bracken BA, Barona A. State of the art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment. *Sch Psychol Int*. 1991;12:119–132. doi:10.1177/0143034391121010

24. Nierenberg AA, DeCecco LM. Definitions of antidepressant treatment response, remission, nonresponse, partial response, and other relevant outcomes: a focus on treatment-resistant depression. *J Clin Psychiatry*. 2001;62(Suppl 16):5–9.

25. Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton depression rating scale. *J Affect Disord*. 2013;150(2):384–388. doi:10.1016/j.jad.2013.04.028

26. Dunlop BW, Gray J, Rapaport MH. Transdiagnostic clinical global impression scoring for routine clinical settings. *Behav Sci (Basel)*. 2017;7(3). doi:10.3390/bs7030040

27. Bland JM, Altman DG. Cronbach's alpha. *BMJ*. 1997;314(7080):572. doi:10.1136/bmj.314.7080.572

28. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994;6:284. doi:10.1037/1040-3590.6.4.284

29. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988;240(4857):1285–1293. doi:10.1126/science.3287615

30. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3 (1):32–35. doi:10.1002/1097-0142(1950)3:1<32::AID-CNCR2820 030106>3.0.CO;2-3

31. Hughes G. Youden's index and the weight of evidence revisited. *Methods Inf Med*. 2015;54(6):576–577. doi:10.3414/ME15-04-0007

32. Slocum-Gori SL, Zumbo BD. Assessing the unidimensionality of psychological scales: using multiple criteria from factor analysis. *Soc Indic Res*. 2011;102:443–461. doi:10.1007/s11205-010-9682-8

33. Zimmerman M, Posternak MA, Chelminski I. Using a self-report depression scale to identify remission in depressed outpatients. *Am J Psychiatry*. 2004;161(10):1911–1913. doi:10.1176/ajp.161.10.1911

34. Frank E, Prien RF, Jarrett RB, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry*. 1991;48(9):851–855. doi:10.1001/archpsyc.1991.01810330075011

35. Cheng C, Dong D, He J, Zhong X, Yao S. Psychometric properties of the 10-item Connor-Davidson Resilience Scale (CD-RISC-10) in Chinese undergraduates and depressive patients. *J Affect Disord*. 2020;261:211–220. doi:10.1016/j.jad.2019.10.018

36. Zimmerman M, McGlinchey JB. Depressed patients' acceptability of the use of self-administered scales to measure outcome in clinical practice. *Ann Clin Psychiatry*. 2008;20(3):125–129. doi:10.1080/ 10401230802177680

37. Prien RF, Carpenter LL, Kupfer DJ. The definition and operational criteria for treatment outcome of major depressive disorder. A review of the current research literature. *Arch Gen Psychiatry*. 1991;48 (9):796–800. doi:10.1001/archpsyc.1991.01810330020003

38. Chen S, Chiu H, Xu B, et al. Reliability and validity of the PHQ-9 for screening late-life depression in Chinese primary care. *Int J Geriatr Psychiatry*. 2010;25(11):1127–1133. doi:10.1002/gps.2442