# Measurement-based Treatment of Residual Symptoms Using Clinically Useful Depression Outcome Scale: Korean Validation Study

Sang Won Jeon[1], Changsu Han[1], Young-Hoon Ko[1], Seo Young Yoon[1], Chi-Un Pae[2], Joonho Choi[3], Yong Chon Park[3], Jong-Woo Kim[4], Ho-Kyoung Yoon[1], Seung-Duk Ko[5], Ashwin A. Patkar[6], Mark Zimmerman[7]

[1]Department of Psychiatry, Korea University Ansan Hospital, Korea University College of Medicine, Ansan, [2]Department of Psychiatry, College of Medicine, The Catholic University of Korea, Seoul, [3]Department of Psychiatry, Hanyang University Guri Hospital, Hanyang University College of Medicine, Guri, [4]Department of Psychiatry, Kyung Hee University College of Medicine, Seoul, [5]Department of Health Management, Hyupsung University, Hwaseong, Korea, [6]Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, [7]Department of Psychiatry and Human Behavior, Brown University School of Medicine, Rhode Island Hospital, Providence, RI, USA

**Objective:** This study was aimed at evaluating the diagnostic validity of the Korean version of the Clinically Useful Depression Outcome Scale (CUDOS) with varying follow-up in a typical clinical setting in multiple centers.
**Methods:** In total, 891 psychiatric outpatients were enrolled at the time of their intake appointment. Current diagnostic characteristics were examined using the Structured Clinical Interview for DSM-IV (41% major depressive disorder). The CUDOS was measured and compared with three clinician rating scales and four self-report scales.
**Results:** The CUDOS showed excellent results for internal consistency (Cronbach's α, 0.91), test-retest reliability (patients at intake, $r=0.81$; depressed patients in ongoing treatment, $r=0.89$), and convergent and discriminant validity (measures of depression, $r=0.80$; measures of anxiety and somatization, $r=0.42$). The CUDOS had a high ability to discriminate between different levels of depression severity based on the rating of Clinical Global Impression for depression severity and the diagnostic classification of major depression, minor depression, and non-depression. The ability of the CUDOS to identify patients with major depression was high (area under the receiver operating characteristic curve=0.867). A score of 20 as the optimal cutoff point was suggested when screening for major depression using the CUDOS (sensitivity=89.9%, specificity=69.5%). The CUDOS was sensitive to change after antidepressant treatment: patients with greater improvement showed a greater decrease in CUDOS scores ($p<0.001$).
**Conclusion:** The results of this multi-site outpatient study found that the Korean version of the CUDOS is a very useful measurement for research and for clinical practice.

KEY WORDS: CUDOS; Depression; Residual symptom; Psychometrics; Reliability and validity; Cut-off score.

## INTRODUCTION

Unlike other chronic medical diseases, depression's biological markers have not yet been identified. Thus, in an assessment based on the clinician's observation and the patient's report, a method of accurately measuring the severity and functional activity of depression is needed. Moreover, the ability to represent the progress of depression using a quantifiable score is desirable.[1] However, an objective and quantifiable measurement of depression is difficult to obtain due to the many limitations involved, including human resources, time, and costs in actual clinical practice.[2]

The Clinically Useful Depression Outcome Scale (CUDOS) is a useful tool for screening for depression because it fully covers the Diagnostic and Statistical Manual of Mental Disorders 4th edition (DSM-IV) symptoms of major depressive disorder (MDD) and dysthymic disorder.[3] It also assesses the functional aspects of depression and sensitively evaluates remission of depression and its residual symptoms.[4] The CUDOS is a brief self-administered questionnaire. Thus, it incurs low cost and can be applied simply and quickly in clinical practice. It takes two to three minutes to complete and the completed form can be

scored within 15 seconds. Good reliability and validity of the CUDOS have been demonstrated and it is considered clinically useful.[3]

The CUDOS consists of 18 items: 16 assessing the DSM-IV MDD symptoms, one question on psychosocial impairment in daily activities, and one question on quality of life.[3] In the DSM-IV, the occurrence of depressive symptoms for two weeks is required to diagnose MDD, but the CUDOS uses a one-week duration. The shorter evaluation period is used so the scale can be used to measure outcome on a weekly basis. The items on MDD symptoms evaluate how often a subject has experienced the symptom in the last week using a 5-point Likert scale. Total scores range from 0 to 64.

In antidepressant clinical trials, the progression of depression is often evaluated using a clinician rating measure, such as the Hamilton Depression Rating Scale (HAMD).[5] However, the HAMD takes too long to administer in routine clinical practice. Brief self-report questionnaires are considered desirable to minimize the cost and to evaluate the progress of depression frequently and quickly.[1] In this sense, the CUDOS is useful for both clinicians and patients due to its simplicity, which improves the efficiency of clinical consultations. In terms of simplicity, the Patient Health Questionnaire-9 (PHQ-9), which was developed based on DSM-IV criteria, has been the most useful self-administered scale.[6] However, it has only 10 items compared to the CUDOS' 18, so it yields limited information. For example, sleep pattern is not divided into hypersomnia and insomnia in the PHQ-9, but is addressed in one item. Also, the PHQ evaluates decreased appetite and increased appetite in one item. Thus, important information about the patient's symptoms may be missed, which may affect the treatment plan.

The CUDOS has been translated into several languages, but it has not been translated and validated in Korea. It is important that translated self-rating scales for use in cross-cultural settings, especially in non-Western countries, need to be strictly evaluated before clinical application.[7] In this study, we developed and validated the Korean version of the CUDOS to evaluate its potential for cross-cultural application for Korean subjects and to suggest an optimal cut-off score among clinical outpatient samples.

## METHODS

### Design and Setting

We conducted an observational, prospective study with varying follow-up in a typical clinical setting in multiple centers (two university hospitals and three general psychiatric hospitals) serving various regional communities in South Korea. The study protocol was approved by the institutional review board of each medical center.

### Participants

Psychiatric outpatients who satisfied the selection criteria at each medical center were recruited to be participants in this study. Participants were new patients as the first psychiatric examination, or had not received antidepressant or other psychotropic drug treatment within the past 4 weeks. A total of 952 patients were originally screened. The inclusion criteria were (i) patients aged 19 or older, (ii) patients with cognitive function that enabled them to answer the questions appropriately, and (iii) patients who could read and write Korean. The exclusion criteria were (i) patients with current psychotic or manic symptoms, (ii) patients with underlying medical disease that could affect study evaluation, (iii) patients who had participated in any clinical trial in the last six months, and (iv) bereavement. A total of 61 patients were excluded from this study: 21 had cognitive problems, 10 had medical problems, 8 had current psychotic or manic symptoms, and 22 failed to complete all the scales. Thus, the final sample included 891 subjects.

### Instruments

Depressive symptoms were assessed by the Korean version of the CUDOS, two clinician rating scales (the 17-item HAMD[8] and the Clinical Global Impression for Severity [CGI-S][9]), and two self-report scales (the PHQ-9[6] and the Beck Depression Inventory [BDI][10]). In this study, subjects were rated using the Clinical Global Impression for Depression Severity (CGI-DS). In addition, anxiety symptoms were assessed by a clinician rating scale (the Hamilton Anxiety Rating Scale [HAM-A][11]) and a self-report scale (the Beck Anxiety Inventory [BAI][12]). Somatic symptoms were assessed with a self-report scale, the Patient Health Questionnaire-15 (PHQ-15).[13] All Korean versions of the administered measures were validated in previous studies.[14-18]

### Procedures

Three board-certified psychiatrists and three certified psychologists who were fluent in both English and Korean translated the CUDOS into Korean and back-translated it into English. The validated Korean version of DSM-IV criteria was referenced for the translation.[19] Translation

and back-translation of the CUDOS were repeated after state-of-the-art procedures in cross-cultural assessment were implemented.[20] The process was repeated until the clinicians felt the Korean version was equivalent to the English version and suitable for Korean patients. In this process, the investigators edited the parts that were unclear or could be misunderstood by Korean patients, based on their discussions. The final version was reviewed by a professional translator and scholars of Korean literature and was agreed upon by all the investigators.

The interviews and testing were performed by 13 board-certified psychiatrists. All patients were also interviewed by a trained diagnostic rater who applied the Structured Clinical Interview for DSM-IV (SCID).[21] The subject completed the test before meeting the clinician and all psychiatrists were kept blinded to the subject's response on the measure. All subjects provided written informed consent. Interrater reliability was examined in 16 patients and was satisfactory for all scales (HAMD, $r=0.95, p<0.001$; HAM-A, $r=0.64, p<0.001$; and CGI-DS, $r=0.90, p<0.001$).

The test-retest reliability of the CUDOS was examined in 2 samples. Of the 891 subjects at the time of intake appointment, 88 returned to the study center one week later and completed the CUDOS a second time to exam test-retest reliability. A new patients at intake are highly symptomatic and very sensitive to the treatment over short period of time. Symptom changes might occur for one week interval. In contrast, patients in ongoing treatment are more clinically stable and less sensitive to treatment. Thus, a second sample of 28 was depressed patients in ongoing treatment. They had received treatment for at least 4 months and were clinically stable according to clinician's interview and chart review. They completed the CUDOS at the time of their appointment and were asked to complete it again one week later.

Of the 891 subjects, 61 subjects diagnosed as MDD at baseline completed the CUDOS and were evaluated using the HAMD a second time 8 weeks after antidepressant treatment to investigate sensitivity to symptom changes.

### Data Analyses

Missing data were replaced with the median of the completed data in each item. The number of missing values on each item was less than 0.5%. The internal consistency of the CUDOS was evaluated using Cronbach's α and the item-total correlation. Test-retest reliability and convergent and discriminant validities of the CUDOS compared with other measures were assessed using Pearson's correlation coefficient. The ability of the CUDOS to discriminate between different levels of depression severity was investigated based on the rating of CGI-DS and the diagnostic classification of major depression, minor depression, and non-depression. In turn, an analysis of variance and post-hoc comparison of Tukey's honestly significant difference (HSD) test were used. Sensitivity and specificity were evaluated by assessing the receiver operating characteristic (ROC) curve to obtain the optimal cutoff score when screening for major depression. We compared the ROC curves derived from the CUDOS, PHQ-9, BDI, and HAMD. The pairwise comparison of AUCs between CUDOS and other measures was investigated using method suggested by Hanley and McNeil.[22] Sensitivity to symptom changes after antidepressant treatment were investigated with an analysis of variance and Tukey's HSD test. All analyses were conducted using IBM SPSS Statistics ver. 20.0 for Windows (IBM Co., Armonk, NY, USA). All statistical tests were two-tailed.

## RESULTS

### Sociodemographic Characteristics and DSM−IV Diagnoses

A total of 891 subjects including 373 men (41.9%) and 518 women (58.1%) were analyzed in this study. The mean subject age was 43.6±15.4 (mean ± standard deviation) years (range, 20-76 years). Other demographic characteristics are provided in Table 1. The current DSM-IV Axis I diagnoses of the 891 subjects at their initial appointment are given in Table 2. The most frequent DSM-IV diagnosis was MDD (n=366, 41.1%).

Subjects were classified into major depression, minor depression, and non-depression groups. The major depression group consisted of 386 subjects who were diagnosed with MDD (n=366, 41.1%) or bipolar major depressive episode (n=20, 2.2%). The minor depression group consisted of 106 subjects who were diagnosed with depressive disorder not otherwise specified (n=8, 0.9%) or adjustment disorder with depressed mood (n=28, 3.1%) or dysthymic disorder (n=70, 7.9%) and were not diagnosed with major depression. The non-depression group consisted of 399 subjects who were not diagnosed with either major depression or minor depression. In other words, these subjects had any other DSM-IV diagnoses except for minor depression or major depression.

### Reliability

Cronbach's α was 0.91 ($p<0.001$) at baseline. The item−total correlations ranged from 0.19 to 0.91 (mean=0.67) at

**Table 1.** Sociodemographic characteristics of subjects (n=891)

| Characteristic | Number of subjects (%) |
|---|---|
| Sex | |
| Male | 373 (41.9) |
| Female | 518 (58.1) |
| Age (yr) | |
| 20-29 | 159 (17.8) |
| 30-39 | 167 (18.7) |
| 40-49 | 232 (26.0) |
| 50-59 | 193 (21.7) |
| 60-65 | 93 (10.4) |
| >65 | 47 (5.3) |
| Education (yr) | |
| 0-6 | 328 (36.8) |
| 7-9 | 180 (20.2) |
| 10-12 | 275 (30.9) |
| >13 | 108 (12.1) |
| Marital status | |
| Single | 303 (34.0) |
| Married | 379 (42.5) |
| Divorced | 209 (23.5) |

**Table 2.** Current diagnostic characteristics of the subjects (n=891) by the Structured Clinical Interview for the DSM-IV (SCID)

| DSM-IV diagnosis | Number of subjects (%) |
|---|---|
| Major depressive disorder | 366 (41.1) |
| Bipolar disorder (major depressive episode) | 20 (2.2) |
| Depressive disorder not otherwise specified | 8 (0.9) |
| Adjustment disorder (with depressed mood) | 30 (3.4) |
| Adjustment disorder (without depressed mood) | 13 (1.5) |
| Dysthymic disorder | 73 (8.2) |
| Bipolar disorder (not major depressive episode) | 59 (6.6) |
| Generalized anxiety disorder | 65 (7.3) |
| Panic disorder | 125 (14.0) |
| Social phobia | 64 (7.2) |
| Specific phobia | 8 (0.9) |
| Obsessive compulsive disorder | 60 (6.7) |
| Posttraumatic stress disorder | 63 (7.1) |
| Alcohol abuse/dependence | 60 (6.7) |
| Somatoform disorder | 129 (14.5) |
| Schizophrenia | 4 (0.4) |
| Other psychiatric disorder | 89 (10.0) |

Subjects could be given more than one diagnosis.

baseline. The lowest item-scale correlations were for reverse vegetative symptoms (increased appetite [$r$=0.19] and hypersomnia [$r$=0.24]). The test-retest reliability coefficients were 0.81 ($p < 0.001$) in 88 patients at intake and 0.89 ($p < 0.001$) in 28 depressed patients in ongoing treatment

### Convergent and Discriminant Validity

Table 3 shows significant positive correlations between the CUDOS and other measures (all, $p < 0.001$). The CUDOS was more highly correlated with measures of depression (mean $r$=0.80) than with measures of the other symptom domains (mean $r$=0.42).

### The Ability of the CUDOS to Discriminate between Levels of Severity

The mean CUDOS scores of the major depression group, minor depression group, and non-depression group were 40.5±15.7, 32.3±12.7, 10.5±6.68, respectively. The three group analysis of variance was significant (F=62.5; degree of freedom [$df$]=2,888; $p < 0.001$) and the differences among the three groups were significant using Tukey's HSD test.

The ability of the CUDOS to discriminate among different levels of depression severity was investigated using an analysis of variance based on CGI-DS ratings. The total CUDOS score increased with increases in the CGI score: CGI 1 (n=17), 12.2±7.7; CGI 2 (n=138), 19.6±8.4; CGI 3 (n=168), 27.1±10.9; CGI 4 (n=232), 32.0±13.2; CGI 5 (n=230), 49.5±10.3; CGI 6 (n=98), 53.1±10.3; CGI 7 (n=8), 65.0±21.3. Because the number of subjects who

rated a 1 on the CGI (n=17) or a 7 on the CGI (n=8) were considered to be relatively low, the two lowest CGI rating levels (CGI 1, 2) and the two highest CGI rating levels (CGI 6, 7) were combined. The five-group analysis of variance was significant (F=211.6; $df$=4,886; $p < 0.001$). Tukey's test showed that the differences between each adjacent CGI-DS level were significant except for the comparison between subjects with the CGI scores of 5 and (6 and 7) ($p$=0.06).

### Receiver Operating Characteristic Analysis

The area under the ROC curve (AUC) for the CUDOS was 0.867 (standard error=3.5%, $p < 0.001$) for the diagnosis of major depression. The AUC values indicated that the CUDOS had a significantly high level of discrimination (95% confidence intervals of 0.815-0.934). We selected a score of 20 as the optimal cutoff point when screening for major depression using the CUDOS because this is associated with a sensitivity of 90% for identifying major depression in the current sample (sensitivity 89.9%, specificity 69.5%). A list of sensitivity and specificity pairs for separate scores is provided in Table 4.

In major depression, the AUCs of the CUDOS, PHQ-9, BDI, and HAMD were all satisfactory (Table 5). The pairwise comparison of AUCs between CUDOS and other measures was investigated.[22] There were no statistically significant differences between each AUC.

### Sensitivity to Change in Symptoms after Treatment

Sixty-one patients diagnosed with MDD at baseline

**Table 3.** Correlations between scores on the CUDOS and related measures

| | | Depression | | | | | Anxiety | | Somatic |
|---|---|---|---|---|---|---|---|---|---|
| | | CUDOS | HAMD | CGI-DS | PHQ-9 | BDI | HAM-A | BAI | PHQ-15 |
| Depression | CUDOS | 1.000 | | | | | | | |
| | HAMD | 0.716 | 1.000 | | | | | | |
| | CGI-DS | 0.776 | 0.904 | 1.000 | | | | | |
| | PHQ-9 | 0.835 | 0.728 | 0.738 | 1.000 | | | | |
| | BDI | 0.857 | 0.737 | 0.714 | 0.818 | 1.000 | | | |
| Anxiety | HAM-A | 0.248 | 0.325 | 0.774 | 0.312 | 0.541 | 1 | | |
| | BAI | 0.394 | 0.541 | 0.678 | 0.381 | 0.382 | 0.742 | 1.000 | |
| Somatic | PHQ-15 | 0.635 | 0.665 | 0.621 | 0.645 | 0.578 | 0.564 | 0.417 | 1.000 |

All correlations are significant at $p<0.001$.
CUDOS, Clinically Useful Depression Outcome Scale; HAMD, Hamilton Depression Rating Scale; CGI-DS, Clinical Global Impression for Depression Severity; PHQ-9, Patient Health Questionnaire-9; BDI, Beck Depression Inventory; HAM-A, Hamilton Anxiety Rating Scale; BAI, Beck Anxiety Inventory; PHQ-15, Patient Health Questionnaire-15.

**Table 4.** The sensitivity and specificity of the CUDOS

| Major depression vs. non-major depression | | |
|---|---|---|
| Cutoff score | Sensitivity (%) | Specificity (%) |
| 16.0 | 95.4 | 57.7 |
| 18.0 | 94.1 | 60.5 |
| 19.0 | 92.3 | 66.3 |
| **20.0** | **89.9** | **69.5** |
| 21.0 | 85.2 | 72.4 |
| 22.0 | 83.3 | 77.4 |
| 23.0 | 83.3 | 80.7 |
| 24.0 | 78.8 | 84.5 |
| 25.0 | 77.8 | 86.8 |

Major depression: major depressive disorder or bipolar major depressive episode.

**Table 5.** Pairwise comparison of receiver operating characteristic (ROC) curves

| | Major depression vs. non-major depression | | |
|---|---|---|---|
| | AUC | SE (%) | 95% CI |
| CUDOS | 0.867* | 3.7 | 0.815-0.934 |
| PHQ-9 | 0.871* | 3.7 | 0.817-0.930 |
| BDI | 0.866* | 3.5 | 0.798-0.889 |
| HAMD | 0.939* | 3.7 | 0.815-0.934 |

AUC, area under the ROC curve; SE, standard error; CI, confidence interval; CUDOS, Clinically Useful Depression Outcome Scale; PHQ-9, Patient Health Questionnaire-9; BDI, Beck Depression Inventory; HAMD, Hamilton Depression Rating Scale.
*$p<0.001$.

completed the CUDOS and were evaluated using the HAMD at a second visit 8 weeks after antidepressant treatment. There were no significant differences in the scores of the two scales at baseline between the 61 patients with MDD who attended follow-up and the 305 patients with MDD who did not attend follow up. The correlation of the HAMD to the CUDOS at follow-up were significant ($r$=0.727, $p<0.001$). At follow-up, the CUDOS and the HAMD scores were significantly decreased (CUDOS, 42.1±10.5 vs. 22.4±13.6, $t$=10.08, $p<0.001$; the HAMD, 22.8±6.6 vs. 12.1±7.9, $t$=12.9, $p<0.001$). Subjects were categorized into a remission group (n=30; HAMD $<7$), a responder group (n=17; ≥50% improvement from baseline to follow-up but not in remission), or a non-responder group (n=14). At follow-up, the mean CUDOS scores of the remission group, responder group, and non-responder group were 9.7±6.7, 23.5±7.7, and 33.2±10.6, respectively. The three-group analysis of variance was significant (F=51.9; $df$=2,58; $p<0.001$). The differences among the three groups were significant using Tukey's HSD test.

## DISCUSSION

The results of this study suggest that the Korean version of the CUDOS is a reliable and valid measure of the severity of depressive symptoms. Consistent with the initial validation study of the English language version of the scale,[3] internal consistency was high, all item-scale correlations were significant, the atypical depression symptoms had the lowest item-scale correlations, test-retest reliability was high, and the CUDOS was more highly correlated with other measures of depression than with measures of other symptom domains. Moreover, the ability of the CUDOS to discriminate among different levels of depression severity was significant and the measure was sensitive to change after treatment.

The AUC of 0.867 indicates that the CUDOS has excellent properties for use as a screening instrument in the identification of major depression. A score of 20 as the optimal cutoff point was suggested when screening for major depression. At this cutoff point, the sensitivity was 89.9% and the specificity was 69.5% in our sample. If the instrument is intended for screening, a large number of subjects

should be included in the patient group first and additional evaluations must be conducted to make accurate diagnoses.[23] No special training is required for administering CUDOS, which is also used in primary care settings other than the psychiatric field.[3] Depressed patients who need immediate treatment can be quickly screened using CUDOS before they are referred to psychiatrists. Thus, the sensitivity of the cutoff point must be close to 90%, even if its specificity is sacrificed. There are five categories in the empirically derived range of the depression severity of the CUDOS: scores of 11-20 represent minimal depression; 21-30, mild depression; 31-45, moderate depression; and 46 or higher, severe depression.[3] The cutoff score of 20 suggested in this study belongs to the upper limit of the minimal depression category, which comes immediately before the start of the mild depression category.

The CUDOS was nearly as highly correlated with HAMD and CGI-DS ($r$=0.716-0.776) as with PHQ-9 and BDI ($r$=0.835-0.857). This slightly differential pattern of correlations may be due to method variance of data collection. It is known that the method of data collection (self-reported vs. clinician rated) affects the degree of correlations.[1] The AUC of HAMD (0.939) is slightly greater than the AUC of other self-report measures (CUDOS, PHQ-9, and BDI: 0.866-0.887). For the same reason, this differential pattern of AUC may be because diagnostic assessment was performed using the Structured Clinical Interview.

The percentage of cases of somatoform disorder was relatively high in our sample. Asian patients with MDD often have comorbid somatoform disorder.[24] Many studies have reported that somatic symptoms are core features of depression in East Asian countries, including South Korea.[25] The high prevalence of somatic symptoms is primarily explained by the conventional concept of a disease as projecting emotional conflict onto physical imbalance and by the pattern of expression of personal relationships with physical language rather than exposing emotional suffering.[26] In Confucian culture, it is not socially acceptable to directly express emotions. East Asian patients do not directly complain of symptoms of depression, but instead complain of accompanying symptoms of depression such as somatic symptoms, anxiety, attention, memory disturbance, or hypochondriasis.[27] The relatively high correlation between the CUDOS and the PHQ-15 ($r$=0.635, $p$<0.001) is consistent with the tendency of depressed East Asian patients to somaticize their problems. Future studies should examine whether it would be beneficial to

include somatic symptom items to a the Diagnostic and Statistical Manual of Mental Disorders, fifth edition-based measure, so that it is more appropriate for Asian patients. Alternatively, using the CUDOS along with the PHQ-15[13] or the Symptom Checklist-90 (SCL-90),[28] which focuses on somatic symptom, may help address this problem.

This study focuses on a practical clinical setting and will be more meaningful to clinicians. However, one important limitation of this study is that subjects are restricted to clinical outpatients. This sample may not be fully representative of community-dwelling populations. As a result, the subjects had a high proportion of depression and the sample size of older adults was relatively small because many elderly subjects were excluded based on the exclusion criteria. Different populations do not respond in the same manner to individual items on the CUDOS. For example, the elderly tend not to express their psychological distress or emotions in the same way as young adults.[29] The generalizability to subject with different socio-demographic (e.g., elderly patients) or clinical characteristics (e.g., medical/surgical patient setting) will need to be validated. Secondly, because the sample size rated on CGI 1 or CGI 7 were considered to be relatively small, the five categories of CGI-DS instead of seven were explored to identify ability to discriminate between levels of severity.

In multinational or multicenter clinical trials, the number of which is sharply increasing these days, all participant countries or centers must provide each other with their measurements, which are secured by the reliability and validity of their evaluations of the trial groups. In depression studies such as antidepressant trials, the rating scale is very important, particularly for objectively evaluating the severity of the symptoms or the treatment effects of antidepressant drugs. The Clinician Rating Scale is still widely used for depression studies.[30] However, research and treatments are usually conducted simultaneously in the busy setting of clinical practice. Accordingly, CUDOS is thought to be useful in depression studies because it saves time and cost, and its reliability and validity have been verified. For the same reasons, it is also expected to be useful for large-scale epidemiological studies.

The results of this multi-site outpatient study found that the Korean version of the CUDOS is reliable, valid, and sensitive to change in a Korean outpatient setting. The CUDOS appears to be a very useful measurement for both clinical practice and research, not only to screen for depression, but to also measure the remission of depression and its residual symptoms.

## REFERENCES

1. Zimmerman M, McGlinchey JB. *Depressed patients' acceptability of the use of self-administered scales to measure outcome in clinical practice. Ann Clin Psychiatry 2008;20:125-129.*
2. Keller MB. *Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. JAMA 2003;289:3152-3160.*
3. Zimmerman M, Chelminski I, McGlinchey JB, Posternak MA. *A clinically useful depression outcome scale. Compr Psychiatry 2008;49:131-140.*
4. 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;47:185-188.*
5. Faries D, Herrera J, Rayamajhi J, DeBrota D, Demitrack M, Potter WZ. *The responsiveness of the Hamilton Depression Rating Scale. J Psychiatr Res 2000;34:3-10.*
6. Kroenke K, Spitzer RL, Williams JB. *The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16:606-613.*
7. Chen C, Lee SY, Stevenson HW. *Response style and cross-cultural comparisons of rating scales among East Asian and North American students. Psychol Sci 1995: 6:170-175.*
8. Miller IW, Bishop S, Norman WH, Maddever H. *The modified Hamilton Rating Scale for Depression: reliability and validity. Psychiatry Res 1985;14:131-142.*
9. Haro JM, Kamath SA, Ochoa S, Novick D, Rele K, Fargas A, et al. *The Clinical Global Impression-Schizophrenia scale: a simple instrument to measure the diversity of symptoms present in schizophrenia. Acta Psychiatr Scand Suppl 2003;(416):16-23.*
10. Beck AT, Steer RA, Carbin MG. *Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev 1988;8:77-100.*
11. Hamilton M. *The assessment of anxiety states by rating. Br J Med Psychol 1959;32:50-55.*
12. Beck AT, Epstein N, Brown G, Steer RA. *An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56:893-897.*
13. Kroenke K, Spitzer RL, Williams JB. *The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64:258-266.*
14. Yi JS, Bae SO, Ahn YM, Park DB, Noh KS, Shin HK, et al. *Validity and Reliability of the Korean version of the Hamilton Depression Rating Scale (K-HDRS). J Korean Neuropsychiatr Assoc 2005;44:456-465.*
15. Han C, Jo SA, Kwak JH, Pae CU, Steffens D, Jo I, et al. *Validation of the Patient Health Questionnaire-9 Korean version in the elderly population: the Ansan geriatric study. Compr Psychiatry 2008;49:218-223.*
16. Kim TS, Jeong SH, Kim JB, Lee MS, Kim JM, Yim HW, et al. *The clinical research center for depression study: baseline characteristics of a Korean long-term hospital-based observational collaborative prospective cohort study. Psychiatry Investig 2011;8:1-8.*
17. Yook SP, Kim ZS. *A clinical study on the Korean version of Beck Anxiety Inventory: comparative study of patient and non-patient. Korean J Clin Psychol 1997;16:185-197.*
18. Han C, Pae CU, Patkar AA, Masand PS, Kim KW, Joe SH, et al. *Psychometric properties of the Patient Health Questionnaire-15 (PHQ-15) for measuring the somatic symptoms of psychiatric outpatients. Psychosomatics 2009; 50:580-585.*
19. Cho MJ, Kim JK, Jeon HJ, Suh T, Chung IW, Hong JP, et al. *Lifetime and 12-month prevalence of DSM-IV psychiatric disorders among Korean adults. J Nerv Ment Dis 2007;195:203-210.*
20. 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.*
21. Hahn OS, Ahn JH, Song SH, Cho MJ, Kim JK, Bae JN, et al. *Development of Korean version of structured clinical interview schedule for DSM-IV axis I disorder: interrater reliability. J Korean Neuropsychiatr Assoc 2000;39:362-372.*
22. Hanley JA, McNeil BJ. *A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839-843.*
23. Nease DE Jr, Maloin JM. *Depression screening: a practical strategy. J Fam Pract 2003;52:118-124.*
24. Ryder AG, Chentsova-Dutton YE. *Depression in cultural context: "Chinese somatization," revisited. Psychiatr Clin North Am 2012;35:15-36.*
25. Zhou X, Min S, Sun J, Kim SJ, Ahn JS, Peng Y, et al. *Extending a structural model of somatization to South Koreans: Cultural values, somatization tendency, and the presentation of depressive symptoms. J Affect Disord 2015; 176:151-154.*
26. Cheng SK. *Understanding the culture and behaviour of East Asians--a Confucian perspective. Aust N Z J Psychiatry 1990;24:510-515.*
27. Pang KY. *Symptoms of depression in elderly Korean immigrants: Narration and the healing process. Cult Med Psychiatry 1998;22:93-122.*
28. Derogatis LR, Unger R. *Symptom checklist-90-revised. In: Weiner IB, Craighead WE, editors. Corsini encyclopedia of psychology. 4th ed. New Jersey:John Wiley & Sons;2010.*
29. Gottfries CG. *Is there a difference between elderly and younger patients with regard to the symptomatology and aetiology of depression? Int Clin Psychopharmacol 1998;13 Suppl 5:S13-S18.*
30. Khan A, Khan SR, Shankles EB, Polissar NL. *Relative sensitivity of the Montgomery-Asberg Depression Rating Scale, the Hamilton Depression rating scale and the Clinical Global Impressions rating scale in antidepressant clinical trials. Int Clin Psychopharmacol 2002;17:281-285.*