



Psychometric Properties of Persian Version of Structured Clinical Interview for DSM-5-Research Version (SCID-5-RV): A Diagnostic Accuracy Study

Parvaneh Mohammadkhani¹, Ameneh Setareh Forouzan², Zahra Hooshyari³ and Imaneh Abasi^{4,*}

¹Social Welfare Management Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

²Social Determinants of Health Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

³Psychiatry and Psychology Research Center, Tehran University of Medical Sciences, Tehran, Iran

⁴Department of Clinical Psychology, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

*Corresponding author: Department of Clinical Psychology, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: abasi@sbmu.ac.ir

Received 2020 January 16; Revised 2020 April 28; Accepted 2020 May 29.

Abstract

Background: The structured clinical interview for DSM-5 has recently been revised to reflect the new findings in the diagnostic criteria of psychological disorders.

Objectives: The present study aimed to evaluate the reliability and validity of the Persian translation of structured clinical interview for diagnostic and statistical manual of mental disorders-fifth edition (DSM-5)-research version (SCID-5-RV) on Iranian adult population.

Methods: In the current diagnostic accuracy study a total of 305 clinical samples were admitted to fifteen adult clinical settings and a subsample of these participants (n = 50, with a mean age of 34.31 and a standard deviation of 11.96) was recruited to evaluate test-retest reliability, and 40 non-clinical participants were recruited to examine construct validity. All participants completed the Millon Clinical Multiaxial inventory-III (MCMI-III) and Brief Symptom inventory (BSI).

Results: SCID psychometric properties indicated an acceptable range for internal consistency (0.95 - 0.99), test-retest reliability (0.60 - 0.79), and Kappa reliability (0.57 - 0.72). Further, the agreement between interviewer and psychiatrist diagnoses was assessed using the Kappa index, and the result was satisfactory. The current diagnostic accuracy study used sensitivity and specificity indexes to assess the diagnostic validity of SCID by positive predictive value and also negative predictive value under the "likelihood ratio" domain. Specificity values for most psychiatric disorders were high; the sensitivity values were to somewhat lower. Furthermore, SCID-5-RV categorical diagnoses demonstrated an acceptable construct validity based on the significant differences between the clinical and non-clinical samples in all subscales of BSI except for phobia as well as all clinical subscales of MCMI-III.

Conclusions: In general, the Persian translation of SCID-5-RV represented acceptable reliability and validity for various categorical diagnoses in different clinical settings.

Keywords: Diagnostic Interview, Psychometrics Properties, Assessment, Psychological Disorders

1. Background

The diagnostic and statistical manual of mental disorders (DSM), as a semi-structured interview mainly intended to facilitate communications among professionals. Besides, the evolution of each edition of DSM (DSM-1 to DSM-5) indicates a trend toward addressing the limitations of the previous versions (1). Despite its limitations, DSM is a valuable instrument for both practice and research purposes (2).

To conduct better assessments, clinicians should obtain more in-depth information. Moreover, in-depth in-

formation improves treatment planning and monitoring that in turn, enhances the reliability and validity of the diagnostic nosology. Further, clinicians should know how to increase clinical utility while addressing the shortcomings of previous versions of the DSM, including high comorbidity rates, the exclusive emphasis on categorical classification, a need for incorporating a dimensional approach, and the excessive use of "not otherwise specified" (3-5). Accordingly, for more than a decade, the DSM workgroup was developing the DSM-5. To this aim, evidence from clinical practice, neurological, genetic literature, epidemiological findings, cognitive neuroscience, and pathophysiol-

ogy, which yielded important insights about pathology, particularly treat debilitation cases, were used (1).

In line with the goal of Research Domain Criteria launched by the National Institute of Mental Health and to address the above-mentioned limitations, the following items were added to the DSM-5: possible shared etiologies, subthreshold symptoms, and cross-cutting symptoms. Besides, by adding specifiers for “not otherwise specified” (NOS) and other categories, supplementing the DSM categorical system with dimensions, and expansion of categories, its clinical utility was expanded (1, 3).

The workgroup who proposed the diagnostic criteria confirmed the face and construct validity of the DSM-5 (6). To assess the reliability of the DSM-5, a selection of DSM-5 categorical diagnoses was used in a field trial, and it was found that most diagnoses with an adequate sample size had a good to very good reliability. Therefore, since DSM is the most widely used interview structure in adult psychopathology and availability of several revisions of DSM-5, assessing the reliability and validity of the DSM-5 diagnosis is of crucial importance.

Semi-structured interviews, such as the SCID, have several advantages over paper-pencil questionnaires. For instance, probing and clarifying the questions and their answers cannot be distorted by the respondents' reading level, which allows for assessing a wide area of psychopathology categorically or dimensionally and accurately capturing the multifactorial nature of psychopathology. The structural clinical interview is used to determine DSM-5 axis I disorders, and it is also useful for research purposes (7, 8).

To the best of our knowledge and for linguistic adaptations, this is the first study that translated SCID5-RV in Iran by implementing it in a large and diverse adult clinical sample and assessed the reliability and validity of the SCID-5-RV. Accordingly, the psychometric properties of the latest version of SCID were evaluated to see whether it has adequate reliability and validity in the nonwestern country. Additionally, the current study paves the way for future research to evaluate DSM-5 criteria on the Iranian population or for applying it in other studies and clinical practices.

2. Objectives

Thus, the current study aimed to assess the reliability of the Persian version of SCID-5-RV, including test-retest reliability, and to evaluate the diagnostic accuracy of SCID-5 v, including the criteria validity, sensitivity, and specificity. Besides, construct validity was also assessed.

3. Methods

3.1. Participants

This is a diagnostic accuracy study that aimed to assess the reliability of the Persian version of SCID-5-RV. Participants were selected among clients who had referred with at least one psychiatric complaint to the psychiatric clinics or hospitals (15 facilities for adults in Tehran) during 2018 - 2019. The research proposal, study measures, and informed consent form were approved by the Ethics Committee (no.: IR.USWR.REC.1395.264).

To recruit participants, first, the city of Tehran was divided into northern, southern, eastern, and western districts, and a clinical facility was randomly selected in each district. Finally, based on pre-determined criteria, eligible clinical participants were chosen. Inclusion criteria were as follow: being born in Iran, the ability to speak in Persian, aged more than 18 years, having mental disorders symptoms, and signing the informed consent to participate in the study. The exclusion criteria were having hearing and speech problems, intellectual problems, receiving different treatments, and failure to cooperate with the researcher due to restlessness or severe psychosis. In total, 305 clinical participants were admitted to outpatient or inpatient facilities, and 40 non-clinical participants, mostly staff members of the hospitals, were selected.

While this study is an accuracy study design, the below formula was used to calculate the sample size.

$$n = \left(\frac{Z_{1-\frac{\alpha}{2}}^2 \times SN \times (1 - SN)}{L^2 \times (1 - Prevalence)} \right) \quad (1)$$

Which the anticipated specificity and the value of L were considered as 0.80 and 0.05, respectively. Besides, according to the literature, the prevalence of mental disorders in the general population was estimated at 15%. Placing these values in the above formula resulted in a sample size of 289. However, due to reasons such as failure of samples, we decided to increase the sample size to 370 n (289 + 25%). Eventually, 305 valid clinical participants were eligible for data analysis.

3.2. Instruments

The SCID-5-RV, Millon Clinical Multiaxial inventory-III (MCMI-III), and Brief Symptom inventory (BSI) were used in this study.

The Persian translation of SCID-5-RV was administered, and the procedure was supervised by the first author, who published the Persian version of SCID-I and is one of the

contributor members of reviewing Structured Clinical interview for DSM-5. The SCID-5-RV normally administers in a single session and takes 45 - 90 minutes to administer. To achieve a linguistic adaptation, the following steps were performed: forward translation by bilingual (English/Persian) translators, evaluation, and revision of the translated SCID-5-RV in an expert panel comprised of mental health professionals (three psychiatrists, five clinical psychologists, and three translators). After comparing the English and Persian versions and approving content validity, it was confirmed that both versions are identical, and the only difference about them is the language. Further, to assess the understandability and applicability of the instrument, the Persian version of SCID-5-RV was piloted on a small sample of clinical participants ($n = 40$).

The MCMI-III is a well-known, true-false, and self-administered questionnaire for assessing DSM-based disorders, which contains 175 items about the severity of symptoms and 24 scales (14 for basic personality, and 10 for the clinical disorder) (9, 10). The present study utilized 10 clinical scales, including generalized anxiety, somatic symptom, bipolar disorder, persistent depressive disorder, alcohol and drug use, post-traumatic disorder, thought disorder, major depression, and delusional disorder. It is proved that MCMI-III subscales have very good validity, including high predictive validity (ranged from 0.92 to 0.98), in the Iranian population (11). In the current study, the internal consistency of the clinical scales was in an acceptable range of 0.72 - 0.96.

Brief Symptom inventory (BSI) is a self-administered questionnaire that comprises 35 items derived from the longer 90-item symptom check-list-90 R (SCL-90-R), and each item has a 4-point Likert-type scale ranging from 0 (not at all) to 4 (extremely). BSI assesses a series of psychological symptoms encompassing depression, irritability, phobic anxiety, psychoticism, interpersonal sensitivity, somatization, obsessive-compulsive, anxiety, hostility, and paranoid ideation (12). Internal consistency for all nine subscales of BSI dimensions ranged 0.75 - 0.89, which is acceptable. BSI also showed satisfactory construct validity (13). The internal consistency of BSI subscales in the Iranian population is reported to be 0.71 - 0.96 and yielded satisfactory construct validity (14). In the current study, the internal consistency of BSI subscales was in an acceptable range of 0.88 - 0.96.

3.3. Training

57 psychologists (by local announcements) with a master's degree or higher agreed to contribute as an inter-

viewer in the present study. All volunteers were assessed by the first author to ensure their diagnostic competency as well as interviewing skills, and being well experienced. In total 37 psychologists were selected as an interviewer. Then, two one-day training workshops were held with a 45-day interval. The first workshop was focused on the purpose of the study, the main goals of DSM-5, the structure and content of the SCID-5-RV, the changes made in DSM-5 and SCID-5-RV, and the interviewing skills. After the workshop, each psychologist was asked to interview with two clinical participants and fill in the DSM-5 checklist for SCID-5-RV at the next workshop, followed by checking and evaluating all completed interviews. Finally, 23 psychologists, who had good skills in interviewing, were asked to contribute to the current study as interviewers and were continuously supervised throughout the study implementation.

3.4. Procedure

At each facility, a coordinator (with an MSc degree in clinical psychology) arranged all steps before interviews, including checking the inclusion and exclusion criteria, debriefing participants about the purpose of the study, and obtaining written informed consent. Then, the coordinators were providing the interviewers with sufficient information about the interviewees.

From 370 clinical participants and a total of 45 nonclinical participants, 65 clinical participants and five nonclinical interviews were removed due to incomplete information. Ultimately, MCMI-III and BSI were filled by the participants with different order of measures.

Additionally, to assess the test-retest reliability, 50 interviews with clinical participants were randomly selected, and after two weeks, another interview was performed by an independent interviewer who was not aware of the intention of the interviews.

To study the construct validity, SCID-5-RV diagnoses were compared with the clinical diagnoses made by blinded psychiatrists. For this purpose, all 350 participants were interviewed by a psychiatrist or clinical psychologist, the time limit between 2 measures was 7 days. and the agreement between the SCID-5-RV interview and diagnosis made by psychiatrists were compared.

3.5. Data Analysis

Data were analyzed using SPSS software version 21 (15). Descriptive statistics such as mean, standard deviation, and frequency were estimated for categorical diagnoses (disorder) and demographic information. Regarding the

reliability and validity, the zero-order correlation was used to examine the concordance between diagnosis made by two distinct examiners and the diagnosis based on the SCID-5-RV interview and the psychiatrist's interview, respectively. The following standards were applied to examine the agreement between the diagnosis of the psychiatrist and SCID-5 diagnosis:

Intra-class kappa was considered unacceptable, questionable, good, very good, and excellent if it was below 0.20, 0.20 - 0.39, 0.40 - 0.59, 0.60 - 0.79, and above 0.80, respectively (16).

In addition to examining the internal consistency, Cronbach's alpha was also calculated for each categorical diagnosis. The following rule of thumb was set for an acceptable range:

"> 0.9 = excellent; > 0.8 = good; > 0.7 = acceptable; > 0.6 = questionable; > 0.5 = poor; < 0.5 = unacceptable" (17).

Regarding test-retest reliability, the Pearson coefficient was applied to evaluate the correlation between scores obtained in rounds one and two for each disorder. Further, the following classification was used to confirm the correlation coefficient:

Up to 0.19 = very weak; 0.2 - 0.39 = weak; 0.40 - 0.59 = moderate; 0.6 - 0.79 = strong; 0.8 - 0.9 = very strong; 1.0 = perfect correlation (18).

To assess the validity, other operating characteristics were calculated, including sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio. To obtain the criterion validity, the degree of agreement between the diagnoses made by the assessor and psychiatrists was evaluated using the Kappa coefficient.

To evaluate the construct validity of the SCID-5-RV diagnoses, multivariate analysis of variance (MANOVA) was employed to assess the significant differences between clinical and non-clinical samples in the subscales scores of MCMI-III and BSI. Significance of differences between clinical and non-clinical groups in the pathological scales of MCMI-III and BSI indicated the construct validity of SCID-5-RV diagnoses (19).

4. Results

4.1. Descriptive Findings

The demographic characteristics of the participants and the frequency of their diagnosis are shown in Table 1. As shown in Table 1, of 305 total clinical participants, 169 (55.4%) were female, and their mean age was 34.31 (SD = 11.96).

Table 1. Demographic Characteristics of the Participants and the Frequency of SCID-5-RV Diagnostic Categories^a

Characteristics	Female	Male	Total
Sex	169 (55.4)	136 (44.3)	
Marital status			
Single			152 (49.9)
Married			97 (31.8)
Divorced			27 (8.9)
Widow			29 (9.5)
Education levels			
Under diploma			58 (19)
Diploma			74 (24.3)
Bachelorette			118 (38.7)
MSc			44 (14.4)
Doctoral			11 (3.6)
Job status			
Housekeeper			44 (14)
Unemployed			97 (32)
Employed			164 (54)
Previous treatment			
Yes			25 (9.5)
No			238 (90.5)
Missing			42
Major depressive episode	59 (36.6)	47 (39.8)	106 (38)
Mania	7 (4.3)	13 (11)	20 (7.2)
Hypomania	6 (3.7)	10 (8.5)	16 (5.7)
Cyclothymia	4 (2.5)	8 (6.8)	12 (4.3)
Persistent depressive	19 (11.8)	14 (11.9)	33 (11.8)
Premenstrual dysphoric disorder	16 (10)		
Schizophrenia	8 (5)	13 (11)	21 (7.5)
Schizoaffective	7 (4.3)	13 (11)	20 (7.2)
Delusional disorder	3 (1.9)	3 (2.5)	6 (2.2)
Bipolar-I	10 (6.2)	7 (5.9)	17 (6.1)
Bipolar-II	6 (3.7)	3 (2.5)	9 (3.2)
MDD	35 (21.7)	18 (15.3)	53 (19)
alcohol use	1 (0.6)	9 (7.6)	10 (3.6)
Panic	27 (16.8)	10 (8.5)	37 (13.3)
Agoraphobia	2 (1.2)	4 (3.4)	6 (2.2)
Specific phobia	11 (6.8)	5 (4.2)	16 (5.7)
GAD	40 (24.8)	15 (12.7)	55 (19.7)
Separation anxiety	10 (6.2)	6 (5.1)	16 (5.7)
SAD	17 (10.6)	13 (11)	30 (10.8)
 OCD	24 (14.9)	15 (12.7)	39 (14)
Eating disorder	14 (8.7)	10 (8.5)	24 (8.6)
ADHD	5 (3.1)	6 (5.1)	11 (3.9)
Adjustment Disorder	4 (2.5)	3 (2.5)	7 (2.5)
Acute stress disorder	5 (3.1)	1 (0.8)	6 (2.2)
PTSD	3 (1.9)	1 (0.8)	4 (1.4)

^aValues are expressed as No. (%).

4.2. Reliability

Data related to the reliability of diagnosis are shown in Table 2. Internal consistency of all diagnoses ranged from

$\alpha = .95$ to $\alpha = .99$, which demonstrates an excellent internal reliability. Furthermore, the composite reliability for all diagnoses was $\alpha = .99$, which is highly acceptable (Table 2). Also, the test-retest reliability for all diagnoses ranged from $r = 0.60$ to $r = 0.79$, which indicates that all disorders have good test-retest reliability (Table 2). The Kappa coefficients of two distinct examiners ($n = 65$) ranged from 0.57 to 0.72 (Table 2), which is acceptable.

4.3. Validity

As shown in Table 3, the agreements obtained using κ statistics were good to excellent for all diagnostic categories ($\kappa = 0.63$ to $\kappa = 0.83$). Obsessive-compulsive, eating disorders, and delusional disorders had the largest Kappa coefficient degree, while depressive episode, panic disorder, and acute stress disorder were among disorders with the lowest Kappa coefficient degree.

For most psychiatric disorders, specificity values were high ($n > 0.85$), and the sensitivity values were to somewhat lower (Table 3). In the medical literature, these indexes, which represent the diagnostic validity of the test, are referred to as a form of criterion validity (20). Generalized anxiety disorder and schizophrenia had the highest and the lowest sensitivity, respectively. Schizophrenia, schizoaffective disorder, and bipolar-II had the highest specificity, and panic disorder had the lowest specificity. In addition, LR+ for various categories ranged from 2.26 to 10.89, and LR- of categories ranged from 0.16 to 0.42.

To assess the construct validity of diagnostic categories, clinical and non-clinical participants were compared regarding the subscales of axis-I disorders of MCMI-III and those of BSI by using MANOVA. The multivariate effect was statistically significant for MCMI-III ($F(10, 123) = 3.45, P < 0.001, \text{partial } \eta^2 = 0.22$) and BSI ($F(9, 149) = 3.45, P < 0.02, \text{partial } \eta^2 = 0.12$) subscales. Further, the results of univariate tests indicated significant differences among the groups on all the subscales of MCMI-III, including generalized anxiety, somatic symptom, bipolar disorder, persistent depression disorder, alcohol and drug use, post-traumatic disorder, thought disorder, major depression, and delusional disorder as well as BSI subscales, including somatization, obsession-compulsive, interpersonal sensitivity, depression, anxiety, hostility, paranoid ideation, and psychoticism, except for the phobic anxiety. The results of the univariate analysis of MCMI-III and BSI subscales are provided in Table 4.

5. Discussion

Diagnostic reliability and validity have an important role in the progress of science and the practice of clinical psychology and psychiatry, without which the accurate identification of risk factors, interpretation of psychopathology, and efficacy of treatment would be erroneous (21, 22). Furthermore, cross-cultural and cross-national evaluation of a universal diagnostic interview, such as DSM, are milestones for further studies, particularly clinical trials, in clinical settings.

The findings of the current study are about the initial psychometric properties of SCID-5-RV categorical diagnosis among the Iranian adult population. Generally speaking, SCID-5-RV categorical diagnosis demonstrated good psychometric properties (i.e., excellent internal consistency, test-retest reliability, criterion validity, and acceptable sensitivity, specificity, and construct validity) which are comparable to the properties of the original SCID-5-RV categorical diagnosis (6, 22).

Investigating the reliability showed satisfactory internal consistency, test-retest reliability, and agreement between two distinct examiners. It worth noting that, in the current study, intra-class kappa coefficient and test-retest reliability of diagnosis categories were higher than those in the previous field trial on categorical disorders that used DSM-5 (22), as well as the DSM-IV among the Iranian population (23). Usually, the test-retest method results in a lower kappa coefficient, as in the current study, compared to the joint or inter-rater method (24). However, the results of the present study are in contrast with such a claim, since the obtained kappa coefficient was extremely higher than that of the studies which used two examiners who observed and rated the same interview independently. This may be related to the larger sample size, and higher homogeneity of the interviewers and clinicians who conducted the interviews (in terms of expertise, age, ethnicity, and other related variables which may affect the intended diagnosis). One of the strengths of the current study was the participation of different clinicians with various clinical disciplines, years of clinical practices, ethnicity/race, age, sex, and other characteristics from all around the country, which enhanced the generalizability of the findings.

Additionally, based on the results of the current study, major depressive disorder, acute stress disorder, and panic disorder had the least Kappa coefficient degree, which is as another important finding of this study that may be related to the lower sample size and comorbidity of these disorders, along with several more severe disorders, including other anxiety or mood disorders, which are common

Table 2. Reliability Indices for SCID-5-RV Based on Diagnostic Categories

Current Diagnostic Categories	Total Numbers (N = 305)	α (N = 305)	Test-Retest Agreement (N = 50)	K (N = 65)
Persistent depressive disorder (PDD)	33	0.99	0.79	0.65
Cyclothymic disorder	12	0.99	0.78	0.66
Schizophrenia	21	0.95	0.60	0.68
Schizoaffective disorder	20	0.99	0.76	0.67
Delusional disorder	6	0.99	-	0.69
Bipolar-I	17	0.95	0.78	0.67
Bipolar-II	9	0.99	-	0.69
Major depressive disorder	53	0.99	0.79	0.58
Alcohol use disorder	10	0.99	0.70	0.62
Panic disorder	37	0.98	0.71	0.57
Agoraphobia	6	0.97	-	0.59
Social anxiety disorder (SAD)	30	0.99	0.69	0.59
Specific phobia	16	0.99	0.65	0.61
Generalized anxiety disorder (GAD)	55	0.98	0.67	0.64
Separation anxiety disorder	16	0.96	0.72	0.63
Obsessive-compulsive disorder	39	0.96	0.78	0.72
ADHD	11	0.97	0.65	0.69
Post-traumatic stress disorder (PTSD)	4	0.98	-	0.67
Adjustment disorders	7	0.95	-	0.59
Acute stress disorder	6	0.97	-	0.58
Sleep disorder	7	0.95	-	0.65
Eating disorder	24	0.96	0.76	0.69

in these diagnosis categories (25-27).

In addition, evaluating the internal consistency of diagnosis categories is an important advantage for the present study, which yielded an excellent alpha coefficient for all disorders, indicating a high correlation between the items of each diagnosis. Reliability coefficients are high and it shows that the designed items measure the target construct with high consistency. Such results are not unexpected for a diagnostic tool, because the items are specifically designed for a specified purpose, and the higher the reliability value, the more reliable the measure (28). Furthermore, the adequate internal consistency was observed even for some disorders that their criteria was changed in DSM-5, including the alcohol use or even several specifiers and MDD with melancholic features. The findings of the current study indicated that the changes from DSM-IV-TR to DSM-5 are acceptable, even in Iranian culture.

To assess the validity of the Persian version of the SCID-5, we compared the results obtained from the questionnaire with clinical diagnoses made consensually by psychiatrists. The results indicated the adequacy of agreement, including acceptable kappa coefficient, high specificity, and almost good sensitivity and likelihood ratio for nearly all diagnoses. The results of the current study are very similar to the findings of studies conducted using pre-

vious versions of SCID on Iranian samples (23).

Sensitivity and Specificity should never be interpreted in isolation as a means for evaluating the clinical utility of a measure (29). The sensitivity and specificity are directly related to the diagnosis of positive and negative cases, respectively. A test with 60% sensitivity, correctly classifies 60% of individuals as patients (true positives), and therefore it is unable to identify the remaining 40% of the patients (false negatives). A test with 60% specificity, correctly identify 60% of people who truly are not sick (true negative), but cannot identify 40% of them (false positive). The likelihood ratio is defined as the likelihood that a patient whose test is positive is really sick compared to patients whose test was negative (30). The observed value of these indexes showed that the SCID is sensitive for positive diagnoses and can identify people without disorder correctly. The observed values of LR showed that these diagnoses are of high accuracy. Besides, there was evidence about the construct validity of SCID-5-RV diagnosis categories, representing significant differences between the clinical and non-clinical population in nearly all subscales of the BSI and MCMI-III, which confirm that the criteria of SCID-5-RV disorders could accurately discriminate clinical and non-clinical populations. The results, which represent the variance in the diagnosis criterion of SCID-5-RV categorical di-

Table 3. Criteria Validity for SCID-5-RV Based on Diagnostic Categories by DSM-V

Current Diagnostic Categories	Total Numbers (N = 305)	K (N = 305)	Sensitivity	Specificity	LR+	LR-
Persistent depressive disorder	33	0.71	0.72	0.79	3.39	0.36
Cyclothymic disorder	12	0.74	0.75	0.90	7.25	0.28
Schizophrenia	21	0.74	0.67	0.93	9.33	0.36
Schizoaffective disorder	20	0.76	0.70	0.93	9.80	0.32
Delusional disorder	6	0.79	-	-	-	-
Bipolar-I	17	0.74	0.80	0.84	4.96	0.24
Bipolar-II	9	0.77	0.78	0.93	10.89	0.24
Major depressive disorder	53	0.65	0.68	0.75	2.73	0.42
Alcohol use disorder	10	0.70	0.85	0.87	6.43	0.16
Panic disorder	37	0.63	0.71	0.68	2.26	0.42
Agoraphobia	6	0.67	-	-	-	-
Social anxiety disorder	30	0.69	0.70	0.81	3.75	0.36
Specific phobia	16	0.69	0.73	0.84	4.55	0.32
Generalized anxiety disorder	55	0.74	0.87	0.85	5.72	0.16
Separation anxiety disorder	16	0.73	0.63	0.89	6.04	0.42
Obsessive-compulsive disorder	39	0.83	0.86	0.79	4.04	0.18
ADHD	11	0.75	0.80	0.90	7.73	0.22
Post-traumatic stress disorder	4	0.76	-	-	-	-
Adjustment disorders	7	0.69	-	-	-	-
Acute stress disorder	6	0.66	-	-	-	-
Sleep disorder	7	0.73	-	-	-	-
Eating disorder	24	0.79	78.95	81.25	4.21	0.26

agnoses, reflect the variance in the underlying construct (31).

The present study has several practical and clinical implications. First, it was conducted across a variety of clinical settings, including academic and private clinical settings, thus a diverse clinical population is captured. Besides, it was conducted by interviewers from various fields. In addition, test-retest and multivariate methodological designs are highly effective methods that their application increases the generalizability of the findings. More importantly, the exclusion criteria were minimal and clinical participants were more similar to natural clinical settings. Also, SCID-5-RV categorical diagnoses are highly reliable and valid when applied to the Iranian population and have good cross-cultural psychometric properties. Thus, these categorical diagnoses can be used in different clinical settings by various interviewers.

Despite having several strengths, the current study has some limitations. First, the sample size was inadequate for all disorders, and calculation of some statistics was

not possible for some disorders. Therefore, further studies should evaluate the cross-validation analysis of more diverse disorders with a larger sample size. Accordingly, specific pilot studies are recommended for each spectrum.

In the present study, first, validity and reliability of the Persian translation of SCID-5-RV were examined, and the construct validity of SCID-5-RV categorical diagnoses was assessed using the MANOVA method. Then, BSI and MCMI-III subscales were compared between clinical and non-clinical populations as an innovative and pioneer perspective to assess the construct validity of DSM-5, which can pave the way for further elaboration of categorical diagnoses. Overall, the SCID-5-RV categorical diagnosis had reliable and valid diagnoses for almost all diagnostic categories in clinical settings in Iran. These results contribute to the growing body of evidence supporting the reliability and validity of SCID-5-RV diagnosis categories and representing the cross-cultural use of the instrument. The authors suggest conducting further studies concerning testing and cross-validating the diagnostic criteria. Consistent

Table 4. Comparison Between Normal Participant and Patient to Test of Construct Validity in MCMI-III Base Rate Score and BSI Subscales by Multivariate Analyses

Dependent Variables	Sum of Squares	df	Mean Square	F	Sig
MCMI-III subscales					
Generalized anxiety	11451.944	1	11451.944	13.874	0.0001
Somatic symptom	10920.911	1	10920.911	18.118	0.0001
Bipolar disorder	4235.778	1	4235.778	6.594	0.011
Persistent depression disorder	7784.965	1	7784.965	14.534	0.0001
Alcohol use	4294.983	1	4294.983	10.693	0.001
Drug use	4532.435	1	4532.435	10.621	0.001
Post-traumatic disorder	8382.005	1	8382.005	12.151	0.001
Thought disorder	9471.438	1	9471.438	22.316	0.0001
Major depression	13410.649	1	13410.649	17.807	0.0001
Delusional disorder	3090.774	1	3090.774	11.913	0.001
BSI subscales					
Somatization	263.945	1	263.945	8.758	0.004
Obsession-compulsive	196.714	1	196.714	7.629	0.006
Interpersonal sensitivity	87.278	1	87.278	5.433	0.021
Depression	510.545	1	510.545	13.384	0.0001
Anxiety	341.932	1	341.932	13.713	0.0001
Hostility	127.636	1	127.636	9.802	0.002
Paranoid ideation	78.490	1	78.490	3.986	0.048
Psychoticism	148.616	1	148.616	8.809	0.003
Phobic anxiety	51.751	1	51.751	3.442	0.065

with the National Institute of Mental Health, Research Domain Criteria project aims to identify the symptomatic and biological dimension of psychopathologies. Future studies are required to investigate cross-cutting dimensional measures of DSM-5 and to build a foundation for incorporating dimensional diagnoses into categorical diagnoses, which help to improve case formulation and treatment plan.

Footnotes

Authors' Contribution: PM, ASF, ZH, and IA assisted with the design and completion of the study. PM, ASF, ZH, and IA conceptualized this study. ZH analyzed the data. PM, IA and ASF provided advice on the acquisition and analysis of the data. IA drafted the manuscript. PM, ASF, ZH, and IA edited the drafted manuscript. PM provided senior supervision. All authors read and revised the whole report.

Conflict of Interests: The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this study.

Ethical Approval: All procedures performed in studies involving human participants were following the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Funding/Support: This work was supported by the Deputy of Research of Welfare and Rehabilitation Sciences [grant number: 96/t/1344].

Informed Consent: Informed consent was obtained from all individual adult participants included in the study.

References

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub; 2013.
2. Sanders JL. A distinct language and a historic pendulum: the evolution of the Diagnostic and Statistical Manual of Mental Disorders. *Archives of psychiatric nursing*. 2011;25(6):394–403. [PubMed: 22114794].
3. First MB. Clinical utility in the revision of the Diagnostic and Statistical Manual of Mental Disorders (DSM). *Professional Psychology: Research and Practice*. 2010;41(6):465. doi: 10.1037/a0021511.
4. First MB. Paradigm shifts and the development of the diagnostic and statistical manual of mental disorders: past experiences and future

- aspirations. *The Canadian Journal of Psychiatry*. 2010;**55**(11):692-700. [PubMed: [21070696](#)].
5. Jones KD. Dimensional and cross-cutting assessment in the DSM-5. *Journal of Counseling & Development*. 2012;**90**(4):481-7. doi: [10.1002/j.1556-6676.2012.00059.x](#).
 6. Clarke DE, Narrow WE, Regier DA, Kuramoto SJ, Kupfer DJ, Kuhl EA, et al. DSM-5 field trials in the United States and Canada, part I: study design, sampling strategy, implementation, and analytic approaches. *American journal of psychiatry*. 2013;**170**(1):43-58. [PubMed: [23111546](#)].
 7. Schinka JA. Further issues in determining the readability of self-report items: Comment on McHugh and Behar (2009). *J Consult Clin Psychol*. 2012;**80**(5):952-5. doi: [10.1037/a0029928](#). [PubMed: [22924333](#)].
 8. First MB, Williams JBW, Karg RS, Spitzer RL. Structured clinical interview for DSM-5—Research version (SCID-5 for DSM-5, research version; SCID-5-RV). Arlington, VA: *American Psychiatric Association*. 2015.
 9. Millon T, Davis RD, Millon C. *MCMI-III Manual*. National Computer Systems; 1997.
 10. Millon T, Davis R, Millon C, Grossman III S. *MCMI-III manual*. 3rd ed. Minneapolis, MN: Pearson Assessments; 2006.
 11. Sharifi AA, Moulavi H, Namdari K. The validity of MCMI-III (Millon, 1994) scales. *Knowl Res Appl Psychol*. 2008;**9**(34):27-38.
 12. Derogatis LR. BSI brief symptom inventory. *Administration, scoring, and procedures manual*. 1993.
 13. Boulet J, Boss MW. Reliability and validity of the Brief Symptom Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*. 1991;**3**(3):433. doi: [10.1037/1040-3590.3.3.433](#).
 14. Mohammadkhani P, Dobson KS, Amiri M, Ghafari FH. Psychometric properties of the Brief Symptom Inventory in a sample of recovered Iranian depressed patients. *International journal of clinical and health psychology*. 2010;**10**(3):541-51.
 15. Spss I. IBM SPSS statistics version 21. Boston, Mass: *International Business Machines Corp*. 2012;**126**.
 16. Kraemer HC, Kupfer DJ, Clarke DE, Narrow WE, Regier DA. DSM-5: how reliable is reliable enough? *American Journal of Psychiatry*. 2012;**169**(1):13-5. [PubMed: [22223009](#)].
 17. George D, Mallery M. *Using SPSS for Windows step by step: a simple guide and reference*. United States: Allyn & Bacon; 2003.
 18. Walker J, Almond P. *Interpreting statistical findings: a guide for health professionals and students*. UK: McGraw-Hill Education; 2010.
 19. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychological bulletin*. 1955;**52**(4):281. doi: [10.1037/h0040957](#). [PubMed: [13245896](#)].
 20. Nunnally JC. An overview of psychological measurement. *Clinical diagnosis of mental disorders*. Germany: Springer; 1978. p. 97-146. doi: [10.1007/978-1-4684-2490-4_4](#).
 21. Chmielewski M, Clark LA, Bagby RM, Watson D. Method matters: Understanding diagnostic reliability in DSM-IV and DSM-5. *Journal of Abnormal Psychology*. 2015;**124**(3):764. doi: [10.1037/abn0000069](#). [PubMed: [26098046](#)].
 22. Regier DA, Narrow WE, Clarke DE, Kraemer HC, Kuramoto SJ, Kuhl EA, et al. DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses. *American journal of psychiatry*. 2013;**170**(1):59-70. doi: [10.1176/appi.ajp.2012.12070999](#). [PubMed: [23111466](#)].
 23. Sharifi V, Assadi SM, Mohammadi MR, Amini H, Kaviani H, Semnani Y, et al. A persian translation of the structured clinical interview for diagnostic and statistical manual of mental disorders: psychometric properties. *Comprehensive psychiatry*. 2009;**50**(1):86-91. doi: [10.1016/j.comppsy.2008.04.004](#). [PubMed: [19059520](#)].
 24. Williams JB, Gibbon M, First MB, Spitzer RL, Davies M, Borus J, et al. The structured clinical interview for DSM-III-R (SCID): II. Multisite test-retest reliability. *Archives of general psychiatry*. 1992;**49**(8):630-6. doi: [10.1001/archpsyc.1992.01820080038006](#). [PubMed: [1637253](#)].
 25. Amerio A, Tonna M, Odone A, Stubbs B, Ghaemi SN. Course of illness in comorbid bipolar disorder and obsessive-compulsive disorder patients. *Asian journal of psychiatry*. 2016;**20**:12-4. doi: [10.1016/j.ajp.2016.01.009](#). [PubMed: [27025465](#)].
 26. Olfson M, Mojtabai R, Merikangas KR, Compton WM, Wang S, Grant BF, et al. Reexamining associations between mania, depression, anxiety and substance use disorders: results from a prospective national cohort. *Molecular psychiatry*. 2017;**22**(2):235. doi: [10.1038/mp.2016.64](#). [PubMed: [27137742](#)].
 27. Meurent AE, Kroll J, Ritz T. Panic disorder comorbidity with medical conditions and treatment implications. *Annual review of clinical psychology*. 2017;**13**:209-40. doi: [10.1146/annurev-clinpsy-021815-093044](#). [PubMed: [28375724](#)].
 28. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*. 2015;**22**(4):195. doi: [10.4103/1117-1936.173959](#). [PubMed: [26776330](#)].
 29. Lange RT, Lippa SM. Sensitivity and specificity should never be interpreted in isolation without consideration of other clinical utility metrics. *The Clinical Neuropsychologist*. 2017;**31**(6-7):1015-28. [PubMed: [28593818](#)].
 30. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia Critical Care & Pain*. 2008;**8**(6):221-3. doi: [10.1093/bjaceaccp/mkn041](#).
 31. Westen D, Rosenthal R. Quantifying construct validity: two simple measures. *Journal of personality and social psychology*. 2003;**84**(3):608. doi: [10.1037/0022-3514.84.3.608](#). [PubMed: [12635920](#)].