

Reliability and Validity Study of the Turkish Version of the Bipolar Spectrum Diagnostic Scale



Bahri İNCE¹, Alparslan CANSIZ², Sevinç ULUSOY³, Kasım Fatih YAVUZ⁴, Erhan KURT⁵,
Kürşat ALTINBAŞ⁶

SUMMARY

Objective: The purpose of this study is to evaluate the reliability and validity of the Turkish Version of the Bipolar Spectrum Diagnostic Scale (BSDS).

Method: The study was carried out with 130 patients diagnosed with bipolar I disorder, 15 patients diagnosed with bipolar II disorder, and 38 patients diagnosed with major depressive disorder attending the outpatient psychiatry departments of the Bakırköy Prof. Dr. Mazhar Osman Training and Research Hospital for Mental Health and Neurological Diseases. The Mood Disorder Questionnaire (MDQ) was used for convergent validity. The internal consistency coefficient, item-total score correlation coefficients, test-retest correlation coefficient, confirmatory factor analysis, correlation with concurrent scale, and ROC curve were statistically calculated.

Results: Confirmatory factor analysis indicated that the 20-item version did not show adequate goodness-of-fit. The item 4 with a relatively low regression weight was removed from the model. For the 19-item revised and corrected model, the observed goodness-of-fit indexes were RMSEA = 0.040, CFI = 0.900, GFI = 0.890, IFI = 0.900 and $\chi^2/df = 1.230$. The internal consistency Cronbach's alpha coefficient was 0.831. The correlation coefficient between the Turkish version of the BSDS and the MDQ was 0.54. The cutoff point of the scale calculated by the ROC analysis was 12 with a sensitivity of 78.6% and a specificity of 86.8%.

Conclusion: The Turkish Version of the BSDS, has been shown to be reliable and valid tool for screening bipolar disorder after removal of the item 4 of the original version of the scale.

Keywords: Bipolar disorder, screening test, reliability, validity

INTRODUCTION

Bipolar disorder (BD) is a chronic mood disorder characterized by recurrent episodes of hypomania, mania, depression, and euthymia. BD often presents with a period of depression; and since the diagnosis cannot be made until the onset of the manic or hypomanic episodes, it may not be correctly diagnosed in the early stage of the disease (Ghaemi et al. 1999). Approximately one-third of the patients with

BD have been misdiagnosed at the first consultation with the delay of the correct diagnoses up to ten years after the onset of the symptoms (Drancourt et al. 2013, Lish et al. 1994).

Self-report scales completed by the patients may increase the diagnostic accuracy of clinicians who have to recognize the clinical symptoms in a limited time. There are several psychometric tools developed to assess the different clinical characteristics of mood disorders. The "Mood Disorders

Received: 19.09.2018, **Accepted:** 03.01.2019, **Available Online Date:** 13.06.2019

¹MD., Bakırköy Professor Dr. Mazhar Osman Training and Research Hospital for Mental Health and Neurological Diseases, Raşit Tahsin Mood Clinic, İstanbul, ²Assist. Prof., Harran University Faculty of Medicine, Department of Psychiatry, Şanlıurfa, ³MD., ³Assoc. Prof., Bakırköy Professor Dr. Mazhar Osman Training and Research Hospital for Mental Health and Neurological Diseases, Department of Psychiatry, İstanbul, ⁴Assoc. Prof., İstanbul Medipol University, Department of Psychology, İstanbul, ⁶Assoc. Prof., Selçuk University, Department of Psychiatry, Mazhar Osman Mood Clinic, Konya, Turkey.

e-mail: kursatalinbas@gmail.com

<https://www.doi.org/10.5080/u23605>

Questionnaire - MDQ” developed by Hirschfeld et al. (2000), the “Hypomania Checklist-32-Revised” developed by Angst et al. (2005), and the “Bipolar Prodrome Symptom Scale” developed by Correll et al. (2014) were adapted to the Turkish language and their reliability and validity were demonstrated, respectively, by Konuk et al. (2007), Vahip et al. (2017) and Aydemir et al. (2018). The BSDS, which assists the screening for both the severe and mild symptoms of BD, was developed by Ronald Pies, and its specificity and sensitivity were determined by Ghaemi et al. (2005). While hypomanic and manic symptoms are screened in the MDQ, both depressive and hypomanic/manic symptoms are measured by the BSDS. The sensitivity of the BSDS for bipolar type I disorder was found to be 75% and the sensitivity for other subtypes in the bipolar spectrum was 79% (Ghaemi et al. 2005). The sensitivity of the Turkish version of the MDQ was found to be 64% at the optimal cut off score of 7. Versions of the BSDS in the Spanish language (Vázquez et al. 2010, Zaratiegui et al. 2011), Chinese language (Chu et al. 2010), Korean language (Wang et al. 2008), and the Persian language (Shabani et al. 2009) were validated and demonstrated to have appropriate psychometric properties.

In this study, reliability and validity of the Turkish form of the Bipolar Spectrum Diagnostic Scale were investigated with the aim to determine whether its psychometric properties were adequate and to calculate the specificity and sensitivity for bipolar disorder.

METHOD

Translation Procedure

Approval was obtained from the research team developing the original form of the scale to study a Turkish version of the form. The translation of the scale was carried out by four psychiatrists. After the translated form had been controlled and accepted, we tested the comprehensibility by asking 20 patients with different levels of education to read the scale; then the patients were requested to judge whether the statements represented their mood and whether the sentences were understandable. In line with the feedback received from all the patients, the text was revised and then back-translated into English by a linguist. This back-translated form was evaluated by the developers of the original scale. The final version of the Turkish form was generated after their approval.

Subjects

The study was conducted between January 1, 2018 and March 31, 2018 in patients diagnosed with bipolar disorder at the Raşit Tahsin Mood Disorders Outpatient Unit of Istanbul Bakırköy Professor Dr. Mazhar Osman Training

and Research Hospital for Mental Health and Neurological Diseases. Patients with bipolar disorder (type-I and type-II) having at least completed primary school and being aged between 18 and 65 years who were in remission were subsequently included in the study. At least two experienced psychiatrists diagnosed the patients according to DSM-IV criteria. Remission criteria were determined as scoring less than 5 points on the Young Mania Rating Scale and less than 7 points on the 17-item Hamilton Depression Rating Scale. Patients with physical or mental limitations at a level hindering them to follow the instructions or patients who met the diagnostic criteria of any type of mental disorder other than bipolar type-I and type-II disorders according to DSM-IV were excluded from the study. In order to evaluate the specificity and sensitivity of the scale, a control group was composed of volunteer participants diagnosed with major depressive disorder (MDD), aged between 18 and 65, who had at least completed primary school. The patients with MDD had been admitted to the general psychiatry outpatient clinic of the same hospital and were diagnosed with major depressive disorder according to DSM-IV criteria by at least two psychiatrists. Those who due to limitations of their mental or physical capacity were unable to follow the guidelines or patients meeting the diagnostic criteria of any psychiatric disorder other than major depressive disorder according to DSM-IV were excluded from the control group. In order to determine the distinctive feature of BSDS more precisely, it was decided to enroll major depression patients rather than healthy controls.

Instruments

The BSDS is a self-report scale consisting of two parts. In the first part of the original English text, each of the 19 sentences based on the symptoms of bipolar disorder, end with a space where the participant puts a check mark if the appropriate expressions are used in reference to the described situation.

One point is scored for each check mark. The second part of the BSDS consists of one simple multiple-choice question asking if the text as a whole fits the participant. According to the degree of fit in the second part, the user adds 6, 4, 2 or 0 points. The total score, ranging from 0 to 25, is obtained by adding up the scores of both parts. In order to facilitate clinical decisions, Ghaemi et al. (2005) proposed that total scores of 20-25, 13-19, 7-12 and 0-6 were indicative of BD with, respectively, high probability, intermediate probability, low probability and very low probability.

In this study, the Turkish version of the MDQ, (Numan et al. 2007) with a cut off value of 6/7, was used for convergent validity analysis of the BSDS. The MDQ is a self-report scale consisting of three parts. The first part of the MDQ screens

for a lifetime history of manic/hypomanic symptoms using 13 yes/no items. The second part queries whether any manic/hypomanic symptoms have been experienced simultaneously. The effect of these symptoms on functionality is assessed by a single question in the third part. Also, in this study, the Turkish versions of the Young Mania Rating Scale (Karadağ et al. 2002) and the Hamilton Depression Rating Scale (Akdemir et al. 1996) were used to assess the participants' remission status.

Procedure

The study was approved by the ethics committee of Bakırköy Professor Dr. Mazhar Osman Training and Research Hospital for Mental Health and Neurological Diseases. Written informed consent was obtained for participating in the use of the BSDS, the MDQ, the YMRS and the HAM-D, from the BD patients and for participating in the use of the BSDS from the MDD patients. The BSDS was retested with an interval of 4 weeks on 26 of the BD patients in order to determine its temporal stability (Bujang 2017).

Statistical Analyses

Confirmatory factor analysis procedure in the AMOS version 23 (Byrne 2010) was used to assess the construct validity of the Turkish version the BSDS, while the other analyses were carried out using the SPSS 20.0 program. Validity of the models are evaluated using indicators of the goodness-of-fit (Brown 2006). The chi-square (χ^2) index was sensitive to sample size. The normed chi-square index was obtained by dividing the chi-square index by degrees of freedom (χ^2/df), which made the index less dependent on the sample size (Hoelter 1983). The other fit indices used consisted of the Comparative Fit Index (CFI) (Bentler 1990), the Goodness of Fit Index (GFI), the Bollen's Incremental Fit Index (IFI), and the Root Mean Square Error of Approximation (RMSEA) (Hu and Bentler 1995). The CFI assumes that all latent variables are uncorrelated (null/independence model) and compares the sample covariance matrix with this null model. The GFI is a measure of fit between the hypothesized model and the observed covariance matrix. The IFI compares chi-square and the degrees of freedom for the model tested to the null model. The RMSEA is a measure of the approximate compatibility in the sample population and uses the square root of the averages. Smaller values for the RMSEA indicate a better model fit. A value of 0.05 or less is indicative of an acceptable model fit (Hu and Bentler 1995). CFI, GFI, and IFI values greater than 0.90 indicate an acceptable goodness of fit (Şimşek 2007). Values of CFI, OFI, IFI > 0.900, $\chi^2/df < 5$, and RMSEA < 0.0854 can be used as acceptable goodness-of-fit criteria (Şimşek 2007).

In the reliability analysis, Cronbach's alpha coefficient for internal consistency and item-total score correlation coefficients were calculated. Furthermore, the test-retest reliability was assessed between the two ratings. In order to carry out the validity analyses, the relationships between the scales were examined by the Pearson correlation analysis. ROC analysis was performed between patients with BD and patients with MDD in order to determine the discriminant validity of the Turkish form of the BSDS. Both cut off points were calculated and the specificity and sensitivity of the scale were obtained.

RESULTS

The study was conducted with 130 patients with BD type I, 15 patients with BD type-II, and 38 patients with MDD. Table 1 shows the sociodemographic and clinical characteristics of the participants. The MDD group had a lower mean age and the difference with the BD group was statistically significant ($t = 3.77$, $p < 0.001$).

Construct Validity Analyses

Confirmatory Factor Analysis was used to test the single-factor model of the original scale. According to the calculated fit indices, it was found that the 20-item version of the BSDS did not produce an adequate fit, and due to the low regression weight ($p > 0.05$), it was decided to remove item 4 (stating: "They often put on weight during these periods.") from the model. Conformity indices revealed high covariance-associated measurement errors between items 8 and 17 and between items 11 and 14, and these errors were corrected

Table 1. Socio-Demographic and Clinical Features of the Participants

	BD n=145	MDD n=38	t/x ²	P
Age (mean±S.D.)	39.6±9.66	33.1±8.4	3.77	<0.001
Gender				
-Female	76(52.4%)	21(55.3%)	0.098	>0.05
-Male	69(47.6%)	17(44.7%)		
Education duration (years)(mean±S.D.)	10.9±3.64	10.7±3.5	0.313	>0.05
Marital status				
-Single	54(37.5%)	13(34.2%)		
-Married	69(47.9%)	23(60.5%)	3.1	>0.05
-Widowed, divorced	21(14.6%)	2(5.3%)		
Diagnosis				
-BD type-I	130(89.7%)			
-BD type-II	15(10.3%)			
BSDS score(mean±S.D.)	15.77±4.62	8.42±3.09		

BD: Bipolar disorder, MDD: Major depressive disorder, BSDS: The Bipolar Spectrum Diagnostic Scale, t: Student t test, x²: Chi-Square, S.D: Standard deviation

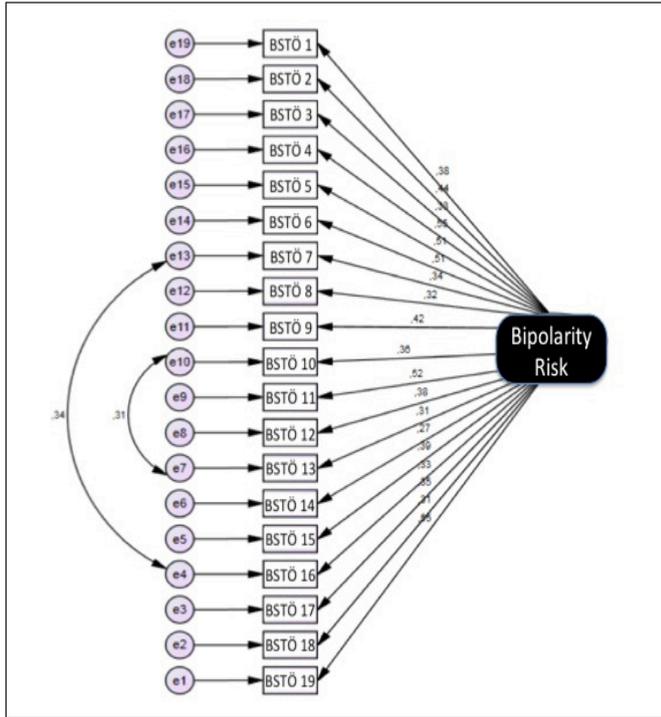


Figure 1. Error Terms and Regression Weights in Revised Single Factor Model of the Bipolar Spectrum Diagnosis Scale

BSTÖ X: Item number of the Turkish version of the BSDS

(Figure 1). According to the final fit indices, the revised and corrected model with 19 items was found to be superior to the previous model (RMSEA = 0.040, CFI = 0.900, GFI = 0.890, IFI = 0.900, and $\chi^2/df = 1.230$) (Table 2). The standardized regression weights for BSDS (0.27 - 0.55, $p < 0.001$) are shown in Table 3.

Concurrent, Convergent and Discriminant Validity Analyses

Concurrent, convergent, and discriminant validity analyses for BSDS were evaluated. Concurrent validity analysis aims to evaluate the correlation of the scale with the expected results. In the Pearson correlation analysis between the total score of the BSDS – following removal of the 4th item – and the total score of the MDQ, a moderate relationship was found ($r = 0.54$, $p < 0.0001$).

Table 3. The Standardized Factor Loads and the Squared Values of Multiple Correlations According to the Confirmatory Factor Analysis for the Single-Factor Structure of the Revised BSDS

Items	Standardized factor loadings	Squared multiple correlations
1	0.384	0.148
2	0.442	0.195
3	0.328	0.107
4	0.551	0.304
5	0.510	0.261
6	0.508	0.258
7	0.340	0.115
8	0.317	0.100
9	0.421	0.177
10	0.358	0.128
11	0.522	0.273
12	0.380	0.144
13	0.311	0.100
14	0.265	0.070
15	0.387	0.150
16	0.326	0.106
17	0.348	0.121
18	0.307	0.095
19	0.545	0.300

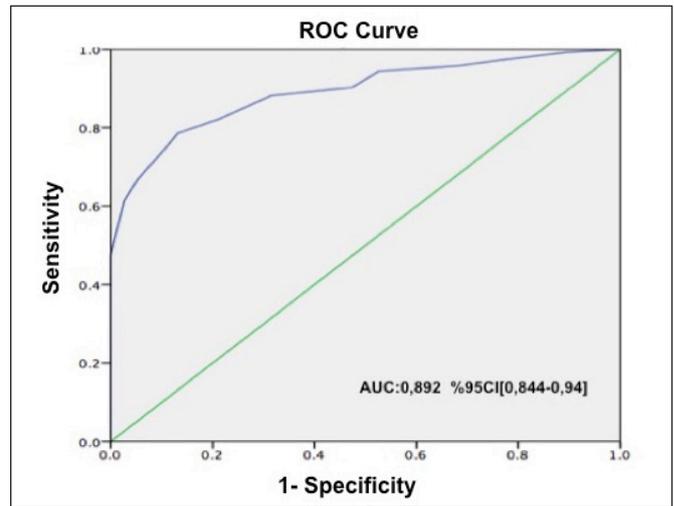


Figure 2. Receiver Operating Characteristic Curve of the Bipolar Spectrum Diagnostic Scale

Table 2. Confirmatory Factor Analysis Results for the BSDS Models

Model	RMSEA	CFI	GFI	IFI	χ^2 / df	p
Single factor model-20 items	0.050	0.814	0.866	0.825	1.355	0.001
Single factor model-19 items (removed 4. item)	0.054	0.807	0.868	0.818	1.413	0.001
Single factor model-19 items (corrected measurement error between items 8 and 17)	0.046	0.856	0.877	0.864	1.310	0.006
Single factor model-19 items (corrected measurement error between items 11 and 14)	0.040	0.900	0.890	0.900	1.230	0.030

RMSEA: Root mean square error of approximation, CFI: Comparative fit index, GFI: Goodness-of-fit index, IFI: Incremental fit index, χ^2/df : The ratio of chi-square fit index to the degree of freedom

After the removal of item 4 in the ROC analysis comparing BD and MDD groups of the Turkish version of BSDS, the area under the ROC curve was 0.892 (Standard error = 0.024), (Figure 2). Using the sensitivity and specificity values of the BSDS, a cut off point of 11/12 was obtained on the ROC curve (Figure 3). For the cut off point of 11, the sensitivity was 82.1% and the specificity was 78.9%; and for the cut off point of 12, the sensitivity was 78.6% and the specificity was 86.8%.

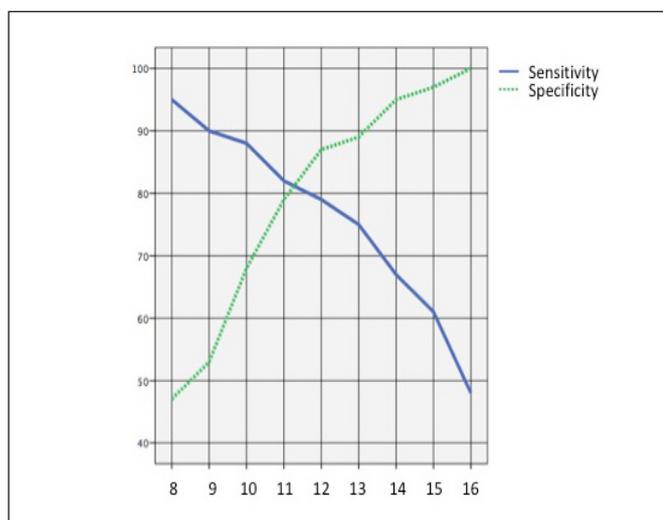


Figure 3. Sensitivity and Specificity of the BSDS Turkish Language Version Across Different Thresholds

Reliability Analyses

Cronbach's alpha coefficient was calculated to determine the internal consistency of the BSDS. After the 4th item was removed in the confirmatory factor analysis, the Cronbach's alpha coefficient was found to be 0.831, indicating a good internal consistency. All items had a positive correlation with the total score of the scale. Table 4 shows the descriptive statistics for items, corrected item-total correlations, and the result of Cronbach's alpha value when items were removed.

In order to determine the temporal stability of the Turkish version of the BSDS, the scale was retested on 26 patients four weeks after the initial test using the test-retest method. The correlation coefficient between the total scores of the test and the retest on the BSDS was calculated to be 0.659 ($p < 0.0001$).

DISCUSSION

Construct Validity Analyses

Confirmatory factor analysis was used on the first version of the BSDS translated to the Turkish language in order to test its construct validity.

On the bases of the obtained goodness-of-fit indices, it was decided to remove item 4 (stating: "They often put on weight during these periods") from the 20-item model due to its low regression weight ($p > 0.05$). All items, other than item 4 about weight gain in depressive episodes, were included in the single-factor structure. The goodness-of-fit indices of the

Table 4. Item-Total Statistics

Items	Scale mean if item deleted	Scale variance if item deleted	Corrected item-total correlation	Alpha if item deleted
1	11.30	19.616	0.329	0.827
2	11.34	18.930	0.484	0.820
3	11.32	20.108	0.176	0.833
4	11.38	19.402	0.333	0.827
5	11.72	19.191	0.354	0.826
6	11.41	19.133	0.390	0.824
7	11.48	18.932	0.418	0.823
8	11.36	19.199	0.399	0.824
9	11.51	18.680	0.472	0.820
10	11.53	18.701	0.463	0.821
11	11.60	18.164	0.590	0.814
12	11.47	18.976	0.409	0.823
13	11.57	18.565	0.491	0.819
14	11.62	19.104	0.361	0.826
15	11.43	18.565	0.527	0.817
16	11.50	18.977	0.400	0.824
17	11.83	19.361	0.352	0.826
18	11.92	19.928	0.253	0.830
19	10.48	16.822	0.575	0.814

revised and corrected 19-item BDS D model was found to be RMSEA = 0.040, CFI = 0.900, GFI = 0.890, IFI = 0.900, and $\chi^2/df = 1.230$ indicating that the construct validity of the revised version of the BDS D was sufficient. In the Chinese language version, item 4 was excluded due to its low factor load. In the Spanish language version, item 4 did not show any difference between BD and MDD (Vázquez et al. 2010). The specificity of this item for the depression episode of BD patients may be low. Atypical antipsychotics have frequently been used in the treatment of mania and maintenance therapy in the recent years (Goodwin et al. 2011). One of the side effects of this class of drugs is conspicuous weight gain, which may occur in non-depressive periods, and item 4 in the original scale may not have been represented in the factor structure.

Reliability Analyses

In the reliability study of the scale, Cronbach's alpha coefficient was found to be as high as 0.831 (Alpar 2013) after exclusion of item 4 according to confirmatory factor analysis. Cronbach's alpha coefficient of the original BDS D was not published and is hence unknown; it was 0.86 for the Persian version and 0.84 for the Spanish one (Zaratiegui et al. 2011). In the Chinese version, intraclass correlation analysis to test the internal consistency revealed a value of 0.81. The Turkish version of the bipolar spectrum diagnostic scale indicated satisfactory internal consistency similar to other studies.

The corrected item-total correlations of the Turkish language version of the BDS D ranged from 0.176 to 0.590 (Table 4). The corrected item-total correlation values for items 3 and 18 of the scale are below 0.3 which is considered to be a sufficient level but not a requirement (Nunnally and Bernstein 1994). If items scoring less than 0.3 for corrected item-total correlations are removed, the risk of narrowing the content of the scale needs to be balanced against the chance of significantly increasing its reliability. When item 3 was removed from the scale, the Cronbach's alpha reliability coefficient increased by 0.02; but this value decreased by 0.01 when item 18 was removed (Table 4). As can be seen, the effect of removing these two items from the scale on Cronbach's alpha reliability coefficient was very low, and therefore, these two items were kept in the scale to prevent the narrowing of the content. Item 3 (stating: "During their 'low' phases, these individuals often feel a lack of energy; a need to stay in bed or get extra sleep; and little or no motivation to do things they need to do.") concurrently measures the lack of energy, increase in sleeping time, and lack of motivation in the depression period. Although these three symptoms are related to one another, they may not always be seen together in the depression period. Hence, asking about these three symptoms jointly may have lowered the item-total correlation value for item 3. Also, presence of atypical depressive episodes and mixed symptoms may be other factors that could affect the value of this correlation. Item 18 (stating: "Sometimes, these

individuals get into difficulty with co- beliefs, workers or the police, during these high periods.") may have been affected by the beliefs and cultural values of individuals in lowering the item's total correlation value.

A moderate level of correlation with a coefficient calculated to be 0.66, was found between the total scores of the BDS D obtained at the initial test and the retest on 26 BD patients four weeks later. The values for this correlation were reported as 0.84 and 0.85, respectively, in the Persian and the Chinese language versions of the BDS D. Compared to the versions adapted to other languages, the temporal stability of the BDS D in our study was found to be low. The higher values of the correlation coefficients in the previous studies may be due to retesting of the scales after 3-5 days and after two weeks in respectively, the Persian and Chinese studies. Also, despite retesting the BDS D in the same euthymic episode, subthreshold symptoms may have differed between the two measurements, which might have caused false negative or positive check markings. Again, factors such as temporary alteration in the physical and mental fatigue status of the participants or the presence of significant stress may have affected the test-retest correlation. Although the test-retest correlation coefficient of the BDS D was moderate in our study, future studies should test the temporal stability of the scale in a shorter time interval with a research design examining the subthreshold symptoms.

Concurrent, Convergent and Discriminant Validity Analyses

When testing the correlation between the total scores of the BDS D and the MDQ, a moderate relationship was found ($r = 0.54$, $p < 0.0001$). Values of this correlation coefficient were 0.63 and 0.582 for, respectively, the Persian and the Korean language versions of BDS D (Bae et al. 2014), and indicate a moderate level of correlation, similarly to our study. Two scales screening for the same psychopathology can be expected to have a higher correlation. However, the total score of the MDQ is only calculated from the first part measuring the manic and hypomanic symptoms. In the BDS D, on the other hand, both the depressive and the manic and hypomanic symptoms are assessed and a total score is obtained by adding a value to each item according to the participant's judgment about its adequacy in describing the differing situations. The diversity and variability in the screened symptoms of the patients may have influenced the correlation between the total scores of the BDS D and the MDQ.

In order to test for criterion validity, the ROC analysis was used to assess the specificity and sensitivity of the BDS D, and the area under the ROC curve (AUC) was found to be 0.892. In the Chinese study, the AUC was determined to be 0.917. Validity of the Turkish language version of the scale was satisfactory regarding the AUC. The cut off point in the Turkish version of the BDS D was found to be 11/12; with a sensitivity of 82.1% and a specificity of 78.9% for the cut

off point of 11, and a specificity of 86.8% and sensitivity of 78.6% for the cut off point of 12 (Figure 3). In the original study of the scale, the cut off score was 13, and sensitivity and specificity for all subtypes of bipolar disorder were 75% and 93%, respectively (Ghaemi et al. 2005). The cut off scores were 12 with 74% sensitivity and 97% specificity for the Chinese language version; 14 with 52% sensitivity and 79% specificity for the Persian language version, 12 with 76% sensitivity and 81% specificity for the Spanish language version and it was 10 with 73% sensitivity and 85% specificity for the Korean language version, indicating that the cut off scores were in the 10-14 range with high specificity and sensitivity levels in the previous studies as well. Given the exclusion of one item from the Turkish language version of the BSDS, a cut off score of 12 was assumed to be appropriate. The sensitivity and specificity of the Turkish version of the MDQ (Numan et al. 2005) were found to be 64% and 77%, respectively, which means that the sensitivity and specificity of the language adapted BSDS versions are higher than those of the MDQ versions.

When evaluating the results of this study, it is useful to consider the limitations. In our study, the incidence of BD type I disorder was 89.7%. The specificity of the original BSDS scale was found to be higher for subtypes of bipolar disorder other than type I. In our study, since the majority of the patients in the BD group were of type I disorder, sensitivity and specificity of the Turkish version of the BSDS were not examined on all BD subtypes. Other limitations that may have affected the findings of our study were that the study did not include a healthy control group, the MDD group on average consisted of younger participants, and the test-retest correlation of the scale was established after more than two weeks.

CONCLUSIONS

The results of this study, aiming to show the reliability and validity of the Turkish version of the BSDS as a screening test that can help to reach an early diagnosis of bipolar disorder, demonstrate that the scale has psychometric properties suitable for use both in clinical practice and in clinical trials.

REFERENCES

- Akdemir A, Dönbak Örsel S, Dağ İ et al (1996) Hamilton Depresyon Derecelendirme Ölçeği (HDDÖ)'nin geçerliliği, güvenilirliği ve klinikte kullanımı. *Psikiyatri Psikoloji Psikofarmakoloji Dergisi*: 3P 4: 251-9.
- Alpar RC (2013) *Uygulamalı çok değişkenli istatistiksel yöntemler*. 4thEd, Ankara, Detay Yayıncılık, p. 849.
- Angst J, Adolfsson R, Benazzi F et al (2005) The HCL-32: towards a self-assessment tool for hypomanic symptoms in outpatients. *J Affect Disord* 88: 217-33.
- Aydemir Ö, Öztekin S, Akdeniz F (2018) Reliability and Validity Study of the Turkish Version of Bipolar Prodrome Symptom Scale. *Turk Psikiyatr Derg* 29: 116-21.
- Bae SO, Kim MD, Lee JG et al (2014) Is it useful to use the Korean version of the mood disorder questionnaire for assessing bipolar spectrum disorder among Korean college students? *Asia Pac Psychiatry* 6: 170-8.
- Bentler PM (1990) Comparative fit indexes in structural models. *Psychol Bull* 107: 238-46.
- Brown TA (2006) *Confirmatory Factor Analysis for Applied Research*, 1stEd, New York, Guilford Press p. 86-8.
- Bujang MA, Baharum N (2017) A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Arch Orofac Sci* 12: 1-11.
- Byrne BM (2010) *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*, 2ndEd, New York, Routledge Academic, p. 17-49.
- Chu H, Lin CJ, Chiang KJ et al (2010) Psychometric properties of the Chinese version of the Bipolar Spectrum Diagnostic Scale. *J ClinNurs* 19: 2787-94.
- Correll CU, Olvet DM, Auther AM et al (2014) The Bipolar Prodrome Symptom Interview and Scale-Prospective (BPSS-P): description and validation in a psychiatric sample and healthy controls. *Bipolar Disord* 16: 505-22.
- Drancourt N, Etain B, Lajnef M et al (2013) Duration of untreated bipolar disorder: missed opportunities on the long road to optimal treatment. *Acta Psychiatr Scand* 127: 136-44.
- Ghaemi SN, Sachs GS, Chiou AM et al (1999) Is bipolar disorder still underdiagnosed? Are antidepressants over utilized? *J Affect Disord* 52: 135-44.
- Ghaemi SN, Miller CJ, Berv DA et al (2005) Sensitivity and specificity of a new bipolar spectrum diagnostic scale. *J Affect Disord* 84: 273-7.
- Goodwin FK, Whitham EA, Ghaemi SN (2011) Maintenance treatment study designs in bipolar disorder: do they demonstrate that atypical neuroleptics (antipsychotics) are mood stabilizers? *CNS Drugs* 25: 819-27.
- Hirschfeld RM, Williams JB, Spitzer RL et al (2000) Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. *Am J Psychiatry* 157: 1873-5.
- Hu LT, Bentler P (1995) Evaluating model fit. *Structural Equation Modeling Concepts, Issues, and Applications*, RH Hoyle (Ed) London, Sage, p. 76-99.
- Hoelter JW (1983) The Analysis of Covariance Structures: Goodness-of-Fit Indices. *Sociol Methods Res* 11: 325-44.
- Karadağ F, Oral ET, Aran Yalçın F et al (2001) Young Mani Derecelendirme Ölçeğinin Türkiye'de geçerlik ve güvenilirliği. *Turk Psikiyatr Derg* 13: 107-14.
- Konuk N, Kıran S, Tamam L et al (2007) Validation of the Turkish Version of the Mood Disorder Questionnaire for Screening Bipolar Disorders. *Turk Psikiyatri Derg* 18: 147-54.
- Lish JD, Dime-Meenan S, Whybrow PC et al (1994) The National Depressive and Manic-depressive Association (DMDA) survey of bipolar members. *J Affect Disord* 31: 281-94.
- Nunnally JC, Bernstein IH (1994) *Psychometric Theory*. 3rdEd, New York, McGraw Hill, p:304-5.
- Shabani A, Koohi-Habibi L, Nojomi M et al (2009) The Persian Bipolar Spectrum Diagnostic Scale and mood disorder questionnaire in screening the patients with bipolar disorder. *Arch Iran Med* 12: 41-7.
- Şimşek ÖF (2007) Yapısal eşitlik modellemesine giriş: Temel ilkeler ve LISREL uygulamaları. Ankara, Ekinoks Yayınları, p. 44-9.
- Wang HR, Kim S, Kang SH et al (2008) The Validation Study of the Korean Version of the Bipolar Spectrum Diagnostic Scale. *J Korean Neuropsychiatr Assoc* 47: 533-9. Korean (English abstract)
- Vahip S, Aydemir Ö, Akkaya C et al (2017) Reliability and Validity Study of the Turkish Version of Hypomania Checklist-32-Revised. *Turk Psikiyatri Derg* 28: 117-23.
- Vázquez GH, Romero E, Fabregues F et al (2010) Screening for bipolar disorders in Spanish-speaking populations: sensitivity and specificity of the Bipolar Spectrum Diagnostic Scale-Spanish Version. *Compr Psychiatry* 51: 552-6.
- Zaratiegui RM, Vázquez GH, Lorenzo LS et al (2011) Sensitivity and specificity of the mood disorder questionnaire and the bipolar spectrum diagnostic scale in Argentinean patients with mood disorders. *J Affect Disord* 132: 445-9.