



Validity and reliability of Persian translation of the Borderline Evaluation of Severity over Time (BEST) questionnaire



Mohammad Reza Azizi¹, Homa Mohammadsadeghi², Kaveh Alavi², Maryam Rasoulia², Nazila Karimzad³, Mehrdad Eftekhari Ardebili*²

Received: 7 Oct 2018

Published: 11 Dec 2019

Abstract

Background: Borderline Evaluation of Severity over Time (BEST) is one of the self-reported tools for evaluation of the severity and track the response of treatment of borderline personality disorder. The present study evaluated the validity and reliability of Persian Translation of the Borderline Evaluation of Severity over Time (BEST) Questionnaire and to compare it with a semi-structured clinical interview for DSM-IV axis II (SCID-II).

Methods: The questionnaire was translated into Persian and then, the content and face validities of the questionnaire were determined. The translated BEST questionnaire and SCID-II were conducted on 33 outpatients and 32 hospitalized patients with diagnosis of borderline personality disorder and 30 patients' companions. Forty-five patients completed the questionnaire again in an interval between 7 to 45 days. The data were analyzed using exploratory factor analysis, paired sample t-test, and the Pearson correlation coefficient and Cronbach's α coefficient.

Results: This study evaluates the content validity, face validity, and criterion validity and reliability of the Persian version of the BEST Questionnaire. The mean scores of the BEST questionnaire were 45.6, 39.2, and 24.3 in in-patients, outpatients, and controls, respectively ($p=0.001$). The mean scores of the BEST questionnaire were 43.7 in the first evaluation, and 41.4 in the second one ($r=0.619$, $p<0.001$). Cronbach's α coefficient was 0.761, and it reached to 0.898 after omitting items 13 to 15. The questionnaire had a two-factor structure, including internal turmoil and the disturbance in interpersonal relationships.

Conclusion: The Persian version of the BEST Questionnaire has a high face and content validity, high criterion validity, moderate to high reliability, and an acceptable two-factor structure.

Keywords: Borderline personality disorder (BPD), Borderline Evaluation of Severity over Time (BEST), Structured Clinical Interview for DSM-IV (SCID-II), Validity, Reliability

Conflicts of Interest: None declared

Funding: None

*This work has been published under CC BY-NC-SA 1.0 license.

Copyright© Iran University of Medical Sciences

Cite this article as: Azizi MR, Mohammadsadeghi H, Alavi K, Rasoulia M, Karimzad N, Eftekhari Ardebili M. Validity and reliability of Persian translation of the Borderline Evaluation of Severity over Time (BEST) questionnaire. *Med J Islam Repub Iran.* 2019 (11 Dec);33:133. <https://doi.org/10.34171/mjiri.33.133>

Introduction

Borderline personality disorder (BPD) is a relatively common disorder, and various sources have estimated its

Corresponding author: Dr Mehrdad Eftekhari Ardebili, eftekhariardebili.m@iums.ac.ir

¹ Chamran Hospital, Tehran, Iran

² Mental Health Research Center, Tehran Institute of Psychiatry, Faculty of Behavioral Sciences and Mental Health, Iran University of Medical Sciences, Tehran, Iran

³ Iran Psychiatry Hospital, Tehran, Iran

↑What is "already known" in this topic:

Borderline personality disorder is a common disorder in psychiatry, which there is a number of tools to evaluate its severity. Despite the translation of one of these tools into Persian, it is not used in clinical evaluation. Therefore this research evaluates the "BEST" questionnaire to introduce a usable clinical tool.

→What this article adds:

The findings of this study showed that the "BEST" questionnaire has an appropriate validity and reliability for clinical and research evaluations and can be used for the periodic survey of borderline personality disorder in the clinic and other health centers.

prevalence about 1.6 % among the general population. This disorder is more prevalent in outpatient psychiatric clinics and inpatients (1). For example in a study, the point prevalence of BPD in a general adult outpatient clinic is reported 22.6 percent (2). Socio-cultural factors like social disintegration can increase the prevalence of borderline personality disorder (3). It is more prevalent in young adults; and is associated with a more frequent concurrent psychiatric disorder, drug, and alcohol-related problems, and marital and job problems (4). BPD is also associated with a substantial cost for society (5, 6).

The essential features of BPD are instability of interpersonal relationships, self-image, and affects, and marked impulsivity (1). Borderline personality disorder is generally described as a multidimensional condition; different dimensions have been emphasized as its core. For example, Zanarini has focused on symptoms related to impairment in cognition, affect, Impulsivity, and problems in interpersonal relationships (7). Hurt has emphasized on three dimensions, including identity, affect, and impulsivity (8). In the last two decades of the 20th century, several tools have been developed to evaluate and diagnose BPD. The most important of these tools are Structured Interview for the DSM Personality Disorder (SIDP), Personality Disorder Examination (PDE), Diagnostic Interview for Personality Disorder (DIPD), The Semi-structured Clinical Interview for DSM-IV Axis II (SCID-II) and Personality Disorder Interview (PDI) (9). Since the beginning of the 21st century along with diagnostic tools, some researchers developed some scales to evaluate the severity of this disorder. There are two types of severity evaluation scales. The first type of scales are interview-based, and the presence of a therapist or interviewer is required. The original version of Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) (9) which is Clinician-administered scale, and The Borderline Personality Disorder Severity Index (BPDSI) (10) are two examples of the first type of scales.

The second type of scales is self-reports. This type is very much considered because less time is spent for completion, they can be done quickly and easily and there is no need for the presence of a therapist or a skilled person in order to complete them. On the other hand, the limitations of these tools, such as the lack of active role in the therapist and reduce patient accuracy to answer the questionnaire should be considered. The most important of these tools include the self-reported version of Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) (11), the Short Version of Borderline Symptom List (BSL-23) (12) and the borderline evaluation of severity over time (BEST) (13).

A study on psychometric features and factor validity of the BEST Questionnaire has been conducted in Turkey (14). Three hundred and six participants, including 201 girls and 105 boys, participated in the study. They completed Borderline Personality Questionnaire, Beck Depression Inventory, Personality Belief Questionnaire, and State-Trait Anxiety scales along with the BEST questionnaire. The result of the study was that the Turkish version of BEST is a valid and reliable scale. They also found a

two-factor structure for the scale.

The aim of this study was to evaluate the validity and reliability of the Persian translation of the Borderline Evaluation of Severity over Time (BEST) Questionnaire.

Methods

In this cross-sectional study, the patients diagnosed by at least a psychiatrist as a borderline personality disorder in two centers in Tehran, including Iran Psychiatric Hospital (inpatients and outpatients) and outpatient service of Tehran Psychiatric Institute, were enrolled. The exclusion criteria were having a psychotic or major mood disorder or being influenced by illegal drugs (for example, intoxicated) at the time of the interviews. The study samples were 65 patients (33 outpatients and 32 inpatients) enrolled in the study consecutively when Structured Clinical Interview for DSM-IV axis II disorders (SCID-II) confirmed the diagnosis. The validity and reliability of the Persian translation of SCID-II interview had been demonstrated previously (15). The control group (n=30) was enrolled from the patient's companions. The age range of the study sample was between 18 and 64 years. Demographic variables (age, gender, marital status, and occupation) and past psychiatric history of the patients, included any previous hospitalization and suicide attempts, were recorded.

The BEST Questionnaire contains fifteen items and three subscales; each item is rated from 1 to 5 on a Likert scale. The subscale A has eight items, which assess the thoughts and feelings during the past month. Subscale B has four items, which are about negative behaviors during the past month. Items in subscales A and B are rated based on their severity (1: None/Slight; 5: Extreme). Subscale C has three items, which evaluate positive behaviors. The items in the subscale C are rated based on their frequencies (1: Almost never; 5: Almost always). The total score of the severity of the disorder is the subtraction of the scores of subscale C from the sum of the scores of the subscales A and B. The resulted scores are between -3 to 57. Then a correction factor of 15 is added to change the range to a positive direction. The final range of the scale is 12-72. The score of 12 shows the best and 72 shows the worst condition.

The English version of the BEST Questionnaire was translated into Persian by the first author. The Persian version of the questionnaire was translated back into English by a university professor in the English Department at Shahid-Beheshti University. A draft of the Persian version was prepared after several meetings between the research team. Content and face validities of the questionnaire were confirmed by 14 psychiatrists of academic staff. They were asked to rate the importance, relevance, and clarity of each item. After review by the experts, the average percentage of importance of items was 86%, the relevance of items was 93% and clarity of items was 83%. The final version was confirmed in an expert panel. Face validity of the scale was confirmed in a survey of 7 patients. Pearson correlation coefficient was used to examine the correlation between SCID-II test scores and BEST Questionnaire scores (criterion validity). Cronbach's alpha coefficient

was calculated to determine the internal consistency of the BEST Questionnaire. The reliability of the BEST Questionnaire was determined by the test-retest method. The average score of the BEST Questionnaire was also calculated and compared between groups (discriminant validity). Test-retest reliability was assessed among 45 patients. The time between two replications was between 7 and 42 days, with an average of 11 days.

Data were entered into SPSS 20 software. Quantitative and qualitative variables were presented with mean and standard deviation or frequency and percentage, accordingly. To compare data between groups, we used t-test, analysis of variance and chi square test, as indicated. To assess the reliability of the scale, we tested the change in the scores between test and retest measurements with paired t-test. Then we used Pearson's correlation coefficient to assess the association of scores of two measurements. The Cronbach's α coefficient was calculated. The correlation between BEST and SCID-II scores was also assessed by Pearson's correlation coefficient. To determine the factor structure of the questionnaire, principal component factor analysis was used. Two rotation methods were tested to obtain the best factor analysis. In the orthogonal rotation, the varimax rotation was used, and in the oblique rotation, the direct oblimin rotation was used. Based on the results of the non-rotated model, direct oblimin rotation was chosen. To avoid deletion of the items, the eigenvalue is considered to be above 0.4. But the $CV = \frac{5.152}{\sqrt{n-2}}$ formula was used to select the items that were loaded on both factors. Given the sample size of 95 people, CV was set at 0.534. P values less than 0.05 were considered as statistically significant.

Results

There were no significant differences between the three groups in terms of gender distribution, educational levels, and socioeconomic status (regardless of having a job or being unemployed) (Table 1). The mean age of the two groups of patients was significantly lower than the control group. The percentage of single subjects in the two groups

was higher than the control group.

In the outpatient group, 27 (81.8%) were previously admitted to psychiatric wards. The median frequency of admission was one time (within the range of 1 to 15 times) for outpatients and three times (within the range of 1 to 10 times) for the in-patient group (Table 2).

In the in-patient group, 23 patients (71.9%) and in the outpatient group, 24 patients (72.7%) had at least one history of suicidal attempt. In the comparison group, two patients (6.3%) had a history of suicidal attempt. The number of people who attempted suicide two or three times was 8 (25%), and 13 the subjects in this group (40.3%) had more than three attempts to commit suicide. In the outpatient group the number of people who had one, two or three and more than three attempts of suicide were 9 (27.35%), 3 (9.1 %) and 9 (27.3%), respectively.

The scores of the BEST questionnaire

The mean \pm SD of SCID-II scores were 7.3 \pm 1.6 and 6.4 \pm 1.2 in in-patients and outpatients, respectively ($t=2.761$, $p=0.008$). The mean \pm SD of the BEST scores were 45.6 \pm 10.3 (range: 23-62), 39.2 \pm 10.1 (range: 21-59) and 24.3 \pm 6.8 (range: 13-50) in in-patients, outpatients and control group, respectively ($f=43.05$, $p<0.001$). The mean score of the BEST Questionnaire was significantly lower in controls compared to both patient groups ($p<0.001$) (discriminant validity). The BEST score was also significantly higher in in-patients compared to outpatients ($p=0.039$). There was a significant correlation between BEST Scores and SCID-II scores ($r=0.483$, $p<0.001$) (criterion validity).

Reliability

The mean \pm SD of the BEST scores were 43.7 \pm 10.2 and 41.4 \pm 11.2, respectively, in two measurements of the patients ($t=1.610$, $p=0.114$). There was a significant correlation between test and retest scores ($r=0.619$, $p<0.001$). Cronbach's α coefficient was 0.761. Cronbach's α coefficient was increased to 0.781, 0.770, and 0.775, respectively, if items 13, 14 and 15 were removed. If the last three

Table 1. Demographic characteristics of the participants in the study

	In- patients (n=32)	Outpatients (n=33)	Control (n=30)	Statistic comparison
Age (Mean \pm SD)	28.3 \pm 6.9	27.7 \pm 6.3	43.7 \pm 12.0	F=33.940 p<0.001
Sex n (%)	Male	19 (59.4)	13 (39.4)	$\chi^2=2.888$ p=0.236
	Female	13 (40.6)	20 (60.6)	
Marital status n (%)	Single	18 (56.2)	21 (63.6)	$\chi^2=27.575$ p<0.001
	Married	8 (25.0)	5 (15.2)	
	Divorced or Widow	6 (18.8)	7 (21.2)	
Job Status n (%)	Without income or housewife	32 (68.8)	24 (72.7)	$\chi^2=5.235$ p=0.073
	With income	10 (31.2)	9 (27.3)	
Education Level n (%)	Elementary or guidance School	5 (15.6)	7 (21.2)	$\chi^2=4.167$ p=0.384
	High School or Diploma	22 (68.8)	16 (48.5)	
	College Level	5 (15.6)	10 (30.3)	

Table 2. History of receiving psychiatric and psychotherapy treatments in participants (Frequency and percentage)

	In- patients (n=32)	Outpatients (n=33)	Control (n=30)
Psychotherapy	1 (3.1)	5 (15.2)	-
Medication	8 (25.0)	7 (21.2)	2 (6.7)
Psychotherapy and Medication	23 (71.9)	20 (60.6)	2 (6.7)

items (subscale C) are removed, the Cronbach's α coefficient reaches 0.898.

Factor structure

To determine the factor structure of the questionnaire, only 12 first items were entered. The Kaiser-Mayer-Olkin (KMO) measure for sampling adequacy was 0.886, and the Bartlett sphericity statistic was 535.6 ($p < 0.001$). Two factors were obtained based on an eigenvalue higher than 1.0 indirect oblimin rotation. The first factor with eigenvalue=3.717 accounted for 31% of the variance and the second factor with eigenvalue=3.319 accounted for 27.7% of the variance. In the direct oblimin rotation, if the lower limit of item loading set at 0.534, items 3, 5, and 6 would not include in any factor. However, if the lower limit is reduced to 0.4, all the items would load on one of the factors. Accordingly, items 3, 6, 7, 8, 10, and 11 are loaded on the first factor, and items 1, 2, 4, 5, 9, and 12 on the second factor. All items in the first factor show the internal turmoil and all items in the second factor describes disturbance in interpersonal relationships (Table 3).

Discussion

The findings of this study indicated that the Persian version of the questionnaire was valid and reliable. There was a moderate correlation between BEST and SCID-II scores ($r=0.483$). The questionnaire could differentiate the severity of symptoms between patients and control group ($p=0.001$). Furthermore, the questionnaire could differentiate between the severity of the symptoms of hospitalized patients and outpatients ($p=0.039$). The Cronbach's alpha coefficient in the full questionnaire was 0.761, and it was 0.889 without the last three items, which indicate an excellent internal consistency between the items. Another important finding of this study was the two-factor structure of the questionnaire. In our study, based on the factor analysis, two factors, including internal turmoil and disturbance in interpersonal relationships, were found as effective factors. In the Turkish study, in the analysis of the factor structure, two independent factors were obtained that were based on the structure of the items. In that study,

the first factor consisted of 12 items that related to subscales of negative thoughts, feelings, and behaviors. The second factor consisted of the last three items (positive behaviors subscale). In fact, in the study conducted in Turkey, the questionnaire consists of two distinct sections. The first section contains prognostic signs of borderline personality disorder, and the second section examines the effectiveness of treatment programs. In our study, the factor analysis was performed on the first 12 terms and two factors of internal turbulence and turmoil in interpersonal relationships were obtained. This finding seems to be in good agreement with the range of symptoms of borderline personality disorder in DSM criteria, and it is also consistent with the patient's clinical presentation. Zanarini et al compare the self-report version of Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) with BEST and concludes that in assessing the patient's relationship problems, the emphasis of the self-report version of ZAN-BPD is more on stormy quality of their relationship, while the focus of BEST is on the shift between idealization and devaluation and doesn't reflect turbulent nature of their relationship. That idea is possibly true because there is no direct question about that, but in this study the second factor was disturbances in interpersonal relationships.

A major obstacle in this study was the lack of cooperation of patients to complete the questionnaire again for the evaluation of retest reliability. The same problem reported in the original research (13), so that the interval between completion of the questionnaires was up to three months. The maximum duration in this study was 42 days.

Conclusion

Our research showed that the Persian version of the BEST Questionnaire has high validity and a moderate to high reliability. Furthermore, it seems to have the potential to be used in clinical settings, and future researches.

Acknowledgment

Authors would like to acknowledge Dr. Seyed Vahid Shariat, Dr. Amir Shabani, Dr. Hamidreza Ahmadkhaniha,

Table 3. Factor analysis of the BEST items

Item	Varimax Rotation		Direct oblimin Rotation	
	Factor One	Factor Two	Factor One	Factor Two
Worrying that someone important in your life is exhausted of you or is planning to leave you		0.824		0.877
Major shifts in your opinions about others such as switching from believing someone is a loyal friend or partner to believing the person is untrustworthy and hurtful		0.733		0.789
Extreme changes in how you see yourself. Shifting from feeling confident about who you are to feeling like you are evil or that you don't even exist	0.562		0.494	
Severe mood swings several times a day. Minor events cause major shifts in moods		0.614		0.544
Feeling paranoid or like you are losing touch with reality		0.540		0.463
Feeling angry	0.555		0.496	
Feelings of emptiness	0.734		0.704	
Feeling suicidal	0.839		0.880	
Going to extremes to try to keep someone from leaving you		0.635		0.608
Purposefully doing something to injure yourself or making a suicide attempt	0.883		0.969	
Problems with impulsive behavior	0.595		0.595	
Temper outbursts or problems with anger leading to relationship problems, physical fights, or destruction of property		0.601		0.527

Dr. Amir-Hossein Jalali, Dr. Rooholah Sedigh, Dr. Mohammadreza Shalbafan, Dr. Shabnam Nohehsara, Dr. Raziye Salhiyan, Dr. Fatemeh Hadi, Dr. Leila Kamalzadeh, Dr. Fathemeh Kashani Nasab for their help in validation of the questionnaire.

Authors are also grateful to Dr. Masoud Ahmadzad, Ali Salahi Yekta, and Dr. Hadi Ranjbar for their help in translation and statistical analysis.

Conflict of Interests

The authors declare that they have no competing interests.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). Washington DC: American Psychiatric Pub; 2013.
2. Korzekwa MI, Dell PF, Links PS, Thabane L, Webb SP. Estimating the prevalence of borderline personality disorder in psychiatric outpatients using a two-phase procedure. *Compr Psychiatry*. 2008;49(4):380–6.
3. Paris J. Personality disorders in sociocultural perspective. *J Pers Disord* [Internet]. 1998;12(4):289–301. Available from: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=9891284&retmode=ref&cmd=prlinks>
4. Swartz M, Blazer D, George L, Winfield I. Estimating the Prevalence of Borderline Personality Disorder in the Community. *J Pers Disord*. 1990;4(3):257–72.
5. van Asselt ADI, Dirksen CD, Arntz A, Severens JL. The cost of borderline personality disorder: societal cost of illness in BPD-patients. *Eur Psychiatry*. 2007;22(6):354–61.
6. Wunsch EM, Kliem S, Kröger C. Population-based cost-offset estimation for the treatment of borderline personality disorder: Projected costs in a currently running, ideal health system. *Behav Res Ther*. 2014;60:1–7.
7. Zanarini MC, Gunderson JG, Frankenburg FR, Chauncey DL. Discriminating borderline personality disorder from other axis II disorders. *Am J Psychiatry* [Internet]. 1990;147(2):161–7. Available from: <http://psychiatryonline.org/doi/abs/10.1176/ajp.147.2.161>
8. Hurt SW, Clarkin JF. Borderline Personality Disorder: Prototypic Typology and the Development of Treatment Manuals. *Psychiatr Ann*. 1990;20(1):13-8.
9. Zanarini MC. Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Pers Disord*. 2003;17(3):233-42.
10. Giesen-Bloo JH, Wouters LM, Schouten E, Arntz A. The Borderline Personality Disorder Severity Index-IV: Psychometric evaluation and dimensional structure. *Pers Individ Dif* [Internet]. 2010;49(2):136–41. Available from: <http://dx.doi.org/10.1016/j.paid.2010.03.023>
11. Zanarini MC, Weingeroff J, Frankenburg F, Fitzmaurice G. Development of the Self-report version of the Zanarini Rating Scale for Borderline Personality Disorder. *Personal Ment Health*. 2015;9:243–9.
12. Bohus M, Kleindienst N, Limberger MF, Stieglitz R-D, Domsalla M, Chapman AL, et al. The short version of the Borderline Symptom List (BSL-23): development and initial data on psychometric properties. *Psychopathology*. 2009;42(1):32–9.
13. Pföhl B, Blum N, St. John D, McCormick B, Allen J, Black DW. Reliability and validity of the Borderline Evaluation of Severity Over Time (BEST): A self-rated scale to measure severity and change in persons with borderline personality disorder. *J Pers Disord*. 2009;23(3):281–93.
14. Ceylan V, Kose S, Akin E, Turkcapar MH. Normative data and factorial structure of the Turkish version of the Borderline Personality Questionnaire (Turkish BPQ). *Psychiatry Clin Psychopharmacol*. 2017;27(2):143–51.
15. Bakhtiari M. Surveying mental disorders in patients with body dysmorphic disorder. Iran University of Medical Sciences; 2000.