Reliability and validity of the Farsi version of the standardized assessment of personality-abbreviated scale

Maryam Sepehri¹, Sara Farhang², Habibeh Barzegar³, Hamidreza Shamekhi¹, Ali Fakhari³, Saeed Dastgiri*⁴

¹ Department of Community and Family Medicine, School of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
² Research Center of Psychiatry and Behavioral Sciences AND Department of Psychiatry, School of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
³ Research Center of Psychiatry and Behavioral Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
⁴ Health Services Management Center, Tabriz University of Medical Sciences, Tabriz, Iran

Abstract

Introduction: A short screening tool for high-risk individuals with personality disorder (PD) is useful both for clinicians and researchers. The aim of this study was to assess the validity and reliability of the Farsi version of the Standardized Assessment of Personality-Abbreviated Scale (SAPAS).

Methods: The original English version of the SAPAS questionnaire was translated into Farsi, and then, translated back into English by two professionals. A survey was then conducted using the questionnaire on 150 clients of primary health care centers in Tabriz, Iran. A total of 235 medical students were also studied for the reliability assessment of the questionnaire. The SAPAS was compared to the short form of Minnesota Multiphasic Personality Inventory (MMPI). The data analysis was performed using receiver operating characteristic (ROC) curve technique, operating characteristic for diagnostic efficacy, Cronbach’s alpha, and test-retest for reliability evaluation.

Results: We found an area under the curve (AUC) of 0.566 [95% confidence intervals (CI): 0.455-0.677]; sensitivity of 0.89 and specificity of 0.26 at the cut-off score of 2 and higher. The total Cronbach’s alpha coefficient was 0.38 and Cohen’s kappa ranged between 0.5 and 0.8.

Conclusion: The current study showed that the Farsi version of the SAPAS was relatively less efficient, in term of validity and reliability, in the screening of PD in the population.

Introduction

Personality disorders (PD) are considered as a major public health problem.⁵ PD are of the hardest to treat among psychiatric conditions and are substantially associated with morbidity.⁵,⁶ PD are common and account for a significant burden of public health concerns. Their community prevalence is estimated to be 3% to 10% of the world population.⁴ In a mental health survey on data of 13 high, middle, and low income countries, the World Health Organization (WHO) estimated the prevalence of PD to be 6.1% in the general population.⁵ Their prevalence is higher in clinical samples and varies from 30% in outpatients⁶ and 40% in inpatients⁵ to higher than 70% in prisoners with psychiatric disorders.⁷ PD, as a comorbidity, aggravates therapeutic outcomes in both primary health care and psychiatric care.¹,⁸ In a community setting, individuals with a PD have a higher probability of being unemployed or divorced, and having a comorbid mental health
problem\textsuperscript{9} and general medical conditions, especially increased risk of cardiovascular disease (CVD).\textsuperscript{10} Thus, the evaluation of the status of PD presents valuable information about the health status of community members. However, this evaluation mostly remains at the level of psychiatrists’ clinical judgment\textsuperscript{11} that can be promoted by the use of standardized assessments. These diagnostic assessments are lengthy and require training;\textsuperscript{12} this is a major problem in the collection of large scale population-based data. This might be the essential reason for the disregarding of PD in population-based studies about psychiatric conditions in some countries (including Iran).\textsuperscript{13} However, according to different studies, the prevalence of PD among psychiatric patients in Iran was high.\textsuperscript{13,14} A study carried out in Kermanshah, Iran, estimated a PD prevalence rate of 8.6\% in the general population.\textsuperscript{14} In addition, another study in this city estimated that the prevalence of PD among inpatients of the psychiatric department of Farabi Hospital was 67.5\%.\textsuperscript{15}

Thus, using an efficient screening interview, even as a part of a two-stage procedure for case identification, may be very useful.\textsuperscript{16} The Standardized Assessment of Personality-Abbreviated Scale (SAPAS) is a short and simple interview-administered screening tool for PD. SAPAS was developed from the semi-structured interview of Standardized Assessment of Personality.\textsuperscript{17,18} It was validated in a small sample of psychiatric patients, where it was found to have good psychometric properties, correctly identifying the presence of PD in 90\% of patients, with a sensitivity of 0.94 and specificity of 0.85 at cut-off score of 3.\textsuperscript{12} Moreover, its validity and reliability have been evaluate in various fields including patients with substance use disorder (SUD),\textsuperscript{19,20} depression,\textsuperscript{21} and incarcerated adolescent boys.\textsuperscript{22} SAPAS is short (no interview took longer than 2 minutes to complete), does not require training, is simple to use, and was acceptable to the respondents in the original study.\textsuperscript{12} It, therefore, fulfils many of the criteria for a desirable screening test.\textsuperscript{23} Thus, SAPAS can be used in routine clinical settings for recognizing high-risk individuals, and in epidemiological researches as the first-stage screen for case identification.\textsuperscript{16,24} This paper reports the performance of the Farsi version of SAPAS in a non-clinical sample in Tabriz, in the northwest of Iran.

Methods
This cross-sectional study was performed in two parts. For validity evaluation, 150 individuals were selected from among the clients and staff of primary health care centers in Tabriz. The reliability assessment was conducted on 235 medical students from Tabriz University of Medical Sciences. The exclusion criteria consisted of age of less than 18 years and previously diagnosed psychiatric condition.

SAPAS: The SAPAS consists of eight dichotomously rated items. The validity and reliability of the questionnaire was originally investigated with a sensitivity and specificity of 0.94 and 0.85, respectively, and the positive and negative predictive values of 0.89 and 0.92, respectively.\textsuperscript{12}

Translation and back-translation were used to produce the Farsi version of the SAPAS with authorization from the author. One of the investigators (SF) translated SAPAS into Farsi. The translated version was reviewed by other investigators, and the translator made the necessary modifications. A bilingual psychiatrist (who did not participate in the study) conducted back-translation from Farsi back into English to ensure accuracy.

The results of SAPAS were compared to the short form of Minnesota Multiphasic Personality Inventory (MMPI). This is a standard questionnaire for diagnosing a wide range of self-descriptive characters.\textsuperscript{25}

For validity evaluation, the study was performed in primary health care centers of Tabriz. The participants were interviewed by 10 trained clinical psychologists using the SAPAS. Then, the participants were asked to complete the MMPI. Results were interpreted by one of the investigators (HB).
For reliability assessment, medical students of Tabriz University of Medical Sciences completed the SAPAS after giving informed consent. Retest was performed on the same group 3 to 4 weeks later (25 days on average).

The performance of the test was illustrated by receiver operating characteristic (ROC) curve and the area under the curve (AUC) was calculated. The optimal score was estimated by Youden's index and ranged between 0 and 1; the highest point is usually considered as the best. AUC ranged from 0.50 (in terms of chance) to 1 (perfect performance). AUC values of 0.5-0.7, 0.7-0.9, 0.9-1, respectively, indicate a low, moderate, and high discriminatory ability for the measure. Sensitivity and specificity were also calculated. The reliability was determined by means of both test-retest procedure and Cohen's kappa coefficient calculation. The internal consistency was obtained through Cronbach's alpha.

Results

Validity evaluation: In this phase, 150 subjects with the mean age ± standard deviation (SD) of 37.16 (± 9.45) years participated. Among them, 64.7% were women and 73.3% were married. Based on the results of the MMPI, 30 of the 150 participants (20.0%) were diagnosed with PD. Due to existence of 4 missing data in the SAPAS scale, ROC was calculated with 146 acceptable questionnaires. The AUC was 0.556 [confidence intervals (CI): 0.455-0.677] (Figure 1); this means that a randomly selected individual with PD is 56% more likely to have a higher score of SAPAS than a randomly chosen participant without PD. Thus, it possess a low capability range of diagnostic accuracy.

Then, optimal point was calculated by means of Youden's index (Table 1); 2 was the optimal cut-off score, with sensitivity of 0.90, specificity of 0.28, and correct classification of 40% of individuals. At cut-off scores of 3 and higher there was more balance between the sensitivity and the specificity with 0.65 and 0.44, respectively. Furthermore, sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR−) were calculated at 4 cut-off points which are all presented in table 2.

<table>
<thead>
<tr>
<th>Cut-off score</th>
<th>Youden's Index</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>LR+</th>
<th>LR−</th>
<th>Correct Classification (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>-0.009</td>
<td>0.931</td>
<td>0.059</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>1.5</td>
<td>0.178</td>
<td>0.897</td>
<td>0.282</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>2.5</td>
<td>0.099</td>
<td>0.655</td>
<td>0.444</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49</td>
</tr>
<tr>
<td>3.5</td>
<td>0.036</td>
<td>0.310</td>
<td>0.726</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49</td>
</tr>
<tr>
<td>4.5</td>
<td>0</td>
<td>0.103</td>
<td>0.897</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64</td>
</tr>
<tr>
<td>5.5</td>
<td>0.034</td>
<td>0.069</td>
<td>0.965</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64</td>
</tr>
<tr>
<td>6.5</td>
<td>0.025</td>
<td>0.340</td>
<td>0.991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64</td>
</tr>
<tr>
<td>8.0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64</td>
</tr>
</tbody>
</table>

PPV: Positive predicative value; NPV: Negative predicative value; LR+: Positive likelihood ratio; LR−: Negative likelihood ratio

Figure 1. Receiver operating characteristic (ROC) curve for diagnostic accuracy of the Farsi version of the Standardized Assessment of Personality-Abbreviated Scale (SAPAS) compared to the Minnesota Multiphasic Personality Inventory (MMPI)
Reliability evaluation: In this phase, the subjects included 235 students from the School of Medicine of Tabriz University of Medical Sciences, with the mean age (± SD) of 22.5 (± 3.17). In addition, 58.3% of the participants were women and 12.3% were married.

To assess the internal consistency, Cronbach’s alpha coefficient was used. Total Cronbach’s alpha was 0.377, representing a weak internal consistency. With the elimination of question number 3 (trust in other people), alpha coefficient increased more compared to the elimination of other questions.

Cohen’s kappa coefficients were obtained through test-retest procedure (Table 3). As shown in table 3, Cohen’s kappa coefficients ranged between 0.5 and 0.8, and in a moderate range.

<table>
<thead>
<tr>
<th>Question</th>
<th>Cronbach’s Alpha if item is deleted</th>
<th>Cohen's kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.351</td>
<td>0.784</td>
</tr>
<tr>
<td>2</td>
<td>0.375</td>
<td>0.678</td>
</tr>
<tr>
<td>3</td>
<td>0.417</td>
<td>0.649</td>
</tr>
<tr>
<td>4</td>
<td>0.317</td>
<td>0.715</td>
</tr>
<tr>
<td>5</td>
<td>0.301</td>
<td>0.680</td>
</tr>
<tr>
<td>6</td>
<td>0.309</td>
<td>0.756</td>
</tr>
<tr>
<td>7</td>
<td>0.288</td>
<td>0.740</td>
</tr>
<tr>
<td>8</td>
<td>0.388</td>
<td>0.551</td>
</tr>
</tbody>
</table>

Discussion
The aim of the present study was to evaluate the validity and reliability of the Farsi version of the SAPAS as a screening tool for PD. This study was performed in a non-clinical sample.

The diagnostic efficiency of the Farsi version of the SAPAS was low (AUC < 0.70).26 This rate is comparable with some other studies, including a study performed on the general population,28 and another study on a German outpatient sample.29 However, this rate was higher in the original study introducing the SAPAS,12 a study on a sample of incarcerated adolescent boys,22 and Self-Administered French version of the SAPAS.30 In the current study, the optimal point was proposed to be 2 and higher, by means of Youden’s index. This point was obtained as 3 and higher in the original study on the SAPAS and some other studies. The sensitivity obtained in our study in the score of 2 and higher is comparable with other studies.12,22,30

However, the specificity obtained in our study is lower than that in other studies that may indicate the low diagnosis power of the Farsi version of the SAPAS for non-patients. At the cut-off score of 3, there was a better balance between sensitivity and specificity; however, its diagnostic power was not high. The Cronbach’s alpha calculated in the present study was in the weak range and presented a low internal consistency, which differed from the original study and other clinical studies.12,22,30 Other studies have indicated that the applicability of this tool might also depend on the prevalence of disease, study subjects, and sampling in various settings.29,31

In the present study, internal consistency was lower than some other studies; this is probably due to the low homogeneity among the questions. However, low overall consistency should not be interpreted as an indicator of poor performance of the SAPAS; the content of the SAPAS is multifaceted and this, in turn, is likely to reflect the heterogeneous content of the test concept (i.e., PD).31

Furthermore, Cronbach’s alpha was also influenced by the length of the test. Therefore, if the test is too short, the value of alpha will be reduced.32,33 Our findings are consistent with a study on the validity of three PD screening measures in psychiatric and non-psychiatric samples, in which all three screening scales were shown to be more effective in a psychiatric sample than in a non-psychiatric sample.34 In the non-psychiatric sample, none of the three screens had a statistically significant AUC, and their diagnostic efficiency (percentage of correct classification) was between 50-60%.34 Moreover, one of the problems of exporting screening measures, developed in clinical...
samples, to nonclinical samples may be spectrum bias. A limitation of this study was that the study sample did not reflect the general population and the participants had high educational levels. Therefore, there might be a low prevalence of PD in this population that might affect the results. Moreover, like all cross-cultural studies, the questions or items can have a different meaning in the translated instruments resulting from cultural and linguistic differences. This could have an effect on the validity and reliability of the translated measures.

A limitation of this study was that the study sample did not reflect the general population and the participants had high educational levels. Therefore, there might be a low prevalence of PD in this population that might affect the results. Moreover, like all cross-cultural studies, the questions or items can have a different meaning in the translated instruments resulting from cultural and linguistic differences. This could have an effect on the validity and reliability of the translated measures.

**References**


DOI: 10.1383/psy.4.3.4.62444


