Validity and Reliability of the Iranian Rapid Assessment for Psychiatric Illness Screening Instrument (IRA-PISI) in Primary Health Care

Ali Fakhari 1, Hosein Azizi 2,3, Mohsen Mostofi 3, Sona Sadeghpour 3 and Mostafa Farahbakhsh 1,3,*

1Department of Psychiatry, Tabriz University of Medical Sciences, Tabriz, Iran
2Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
3Research Center of Psychiatry and Behavioral Sciences, Tabriz University of Medical Sciences, Tabriz, Iran
*Corresponding author: Research Center of Psychiatry and Behavioral Sciences, Tabriz University of Medical Sciences, Tabriz, Iran. Email: mfarahbakhsh@gmail.com

Received 2021 February 06; Revised 2022 January 02; Accepted 2022 April 14.

Abstract

Background: An estimated 25% to 30% of patients presenting to family physicians have psychiatric disorders (PDs). Accurate diagnosis of various PDs requires valid and well-designed screening tools and psychiatric interviews. There is no rapid and low-cost tool for PD screening in primary health care (PHC).

Objectives: To determine the validity and psychometric properties of the Iranian Rapid Assessment for Psychiatric Illness Screening Instrument (IRA-PISI) in PHC.

Methods: This psychometric study was performed among a convenience sample of 257 outpatients presenting to the first-line health services in Tabriz. Psychiatric interviews via DSM-5 were used as the gold standard. The validation process included the face validity (literature review and expert comments), content validity, estimating the best cutoff point, and measuring the sensitivity and specificity of the questionnaire compared to the psychiatric interview. Temporal and internal reliabilities were measured by test-retest and Cronbach’s alpha coefficient, respectively.

Results: The final questionnaire consisted of 14 items (with a score range of 14 to 70) after removing seven questions from the initial questionnaire (21 items) during the validation process. The sensitivity and specificity of the questionnaire compared to psychiatric interviews were 83.62% and 75.17%, respectively. The overall relevance and the best cutoff point (score) were 0.91% and 28, respectively. Interclass correlation coefficient (ICC) and Cronbach’s alpha coefficient were 96% and 83%, respectively. The ICC of items/questions ranged from 89.3 to 100.

Conclusions: The IRA-PISI has appropriate validity and reliability for screening common PDs among outpatients presenting to first-line healthcare providers.

Keywords: Mental Disorders, Primary Health Care, Psychometrics, Questionnaire, Screening

1. Background

Psychiatric disorder (PD) is a major public health concern around the world (1). Currently, communities are exposed to a warning increase in the incidence of PDs all over the world, due to various social and political turmoil and the continuous waves of violence (2). It is estimated that 25% to 30% of patients presenting to family physicians have PDs. In 2020, mental disorders were responsible for 20% of disabilities, compared to only 9% in 1990 (3).

In Iran, the incidence rate of PDs is growing for many reasons, including population growth, social changes, the collapse of the family foundation, and economic problems (4). Recently, a systematic review and meta-analysis study from Iran revealed that the prevalence rates of PDs in studies that used screening tools and clinical interviews were 31.03% and 25.42%, respectively, compared to 28.70% by screening and 18.60% by interviews in 2007 (3). Identifying and screening mental disorders in the first-line health care system are unavoidable (3). Accurate diagnosis of various PDs requires valid and well-designed screening tools and clinical psychiatric interviews (5). Clinical interviews require high costs and time. Patient self-report scales do not provide valid and reliable diagnoses and sometimes lead to false-positives and/or false-negative implications (6).

Screening for early case detection and treatment of mental disorders can improve the quality of life, and reduce health system costs (7). Tools stemming from the Patient Health Questionnaire (PHQ) and other valid and reliable instruments had the most application in many fields.
and settings (8). However, these numerous tools could not meet the needs of PHC and primary physicians and need too much time and many psychiatric interviews. This psychometric investigation provides a rapid instrument that first-line healthcare providers can use to screen various PDs in the early stage (9, 10). Screening tools should be well-designed, rapid, concise, and understandable, and accurately identify major and common PDs without distress to patients. Access to such a tool is limited. Moreover, screening tools should be localized; thus, tools used in developed countries cannot identify PDs in low and middle-income countries (11).

The primary health care (PHC) system is an effective and important place for detecting and managing mental health disorders. The PHC services are the most readily available means of health care. With high rates of regular contact, primary care providers play a noticeable role in mental health promotion (12, 13).

2. Objectives

There is no rapid and low-cost screening tool for screening common PDs in PHC. This psychometric study aimed to develop and validate the psychometric properties of the Iranian Rapid Assessment for Psychiatric Illness Screening Instrument (IRA-PISI) in the first-line health care system.

3. Methods

3.1. Study Design

This psychometric study was performed to develop and validate the psychometric properties of the IRA-PISI for screening common PDs in first-line outpatient PHC in Tabriz between 2019 and 2020. The study population included outpatients presenting to family physicians and/or first-line providers in three medical centers, including a community health center, Emam Khomeini Clinic, and a private clinic in Tabriz. The validity and reliability were assessed among outpatients presenting to family physicians in the PHC setting. Considering $P = 0.2$, $\alpha = 0.05$, $d = 0.2P$, and 20% compensation for non-respondents, the sample size was calculated at 257 outpatients. A convenience sampling method was used for selecting the sample. Family physicians are the chief first-line healthcare providers in the Iran health system. They provide mainly therapeutic and preventive healthcare services for a large number of community members.

3.1.1. Eligibility Criteria

Inclusion criteria were age 18 - 60 and informed consent, and exclusion criteria were mental disability, cognitive disorder, and major physical illness such as cancer and chronic hepatitis (13).

3.2. IRA-PISI Development Process

3.2.1. Literature Review

We systematically searched Medline (via Ovid), the Cochrane Library, and Psych Info to find the most relevant records, including common psychiatric disorder screening tools or/questionnaires and their contents. The initial search was performed with the most relevant medical subject heading (MeSH) keywords related to "questionnaire," "psychiatric," and "screening" in the title and/or abstract. The initial search was combined with related terms, including tool, psychological, primary health care, PHC, disorder, diagnosis, case finding, validity, reliability, and family physicians. Boolean operators AND, OR, and NOT were used to combine related words and phrases. Grey literature, such as the WHO or CDC websites, was thoroughly explored. There were no time and language restrictions.

3.2.2. Expert Comments

Twelve psychiatrists were recruited to generate the questionnaire. They had work experience of 16.11 (SD = 6.98) years. Tools/questionnaires and their contents identified in the literature review were discussed in the expert panel (psychiatrists) to generate the questionnaire items. The literature review found instruments, including PRIME-MD, PHQ9 and 15, and K10. A focus group discussion (FGD) was used to generate the questionnaire items. The experts discussed the items and questions of the explored tools, identified relevant items, rated them, removed overlap items, and finalized them to generate the IRA-PIS questionnaire. Common psychiatric disorders included depressive, anxiety, psychotic, obsessive-compulsive, bipolar, and personality disorders. The following principles were considered in the development and scoring steps: (1) each question measures only one variable; (2) the questions should be as simple and short as possible; (3) the questions must be comprehensive; (4) the items use positive sentences; (5) the answers should be on a Likert scale, including always (score = 5), often (score = 4), sometimes (score = 3), rarely (score = 2), and never (score = 1).

3.3. Validity

The first version of the tool was shared among 12 experts to be scored in terms of content validity. The experts were asked to rate instrument items in terms of clarity and relevance to the study constructs on a four-point ordinal
scale (1: not relevant, 2: somewhat relevant, 3: quite relevant, and 4: highly relevant). The number of experts giving a rating of 3 or 4 was divided by the total number of experts to obtain the Content Validity Index (CVI) for relevance and clarity of each item. Furthermore, suggestions from the expert panel and item impact scores were used to examine the instrument’s face validity.

Receiver operating characteristic curve (ROC) was used to measure the best cutoff point of the questionnaire. Structured interviews via DSM-5 (a binary outcome) were conducted by a highly experienced psychiatrist as a gold standard. The structured interviews and questionnaires were used for screening the same sample of 257 outpatients referred to the family physician. Then, the best cutoff point, where it had the most sensitivity and specificity, was measured.

After determining the best cutoff point among 257 outpatients referred to the first-line health services (family physicians), the screened outpatients were divided into a binary group (positive screen and negative screen). Table 1 shows the method for calculating the sensitivity and specificity of the questionnaire. By presenting data in a two-way table, the sensitivity and specificity of the questionnaire compared to the gold standard were measured.

### Table 1. Two-way (2 × 2) to Measure Sensitivity and Specificity of the Questionnaire

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Gold Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Positive</td>
<td>TP</td>
</tr>
<tr>
<td>Negative</td>
<td>FN</td>
</tr>
</tbody>
</table>

Abbreviation: TP, true positive; FP, false positive; FN, false negative; TN, true negative.

### 3.4. Reliability

Test-retest was used to measure the temporal reliability of the questionnaire among 30 outpatients of family physicians in the PHC at a one-week interval. The questionnaires were given to outpatients, and they were instructed how to complete them. There was no time limit for answering the questions. The questionnaires were not gathered until their announcement. They would mark the best option describing them and could ask the examiner any questions. The Cronbach Alpha method measured the internal reliability of IRA-PISI among 30 outpatients presenting to family physicians in PHC.

### 3.5. Ethics Approval

This study was extracted from the thesis of Dr. Mohsen Mostofi, a psychiatry student, and approved by the Ethics Committee of Tabriz University of Medical Sciences (IR.TBZMED.REC.1397.186). Written informed consent was obtained before the interviews.

### 3.6. Statistical Analysis

We used SPSS software (version 19.0, Chicago, IL, USA) for data analysis. The Kolmogorov-Smirnov test assessed the data normality. The ROC curve determined the best cutoff point. Screening data are presented in a two-way table in which columns show the number of positive and negative cases by the gold standard method (interviews) and rows show the number of positive and negative cases through questionnaires. Conditional cells were used to calculate the sensitivity and specificity of the questionnaire. Interclass correlation coefficient (ICC) and Cronbach’s alpha coefficient were used to measure the temporal and internal reliability of the questions/items among 30 outpatients at a one-week interval, respectively.

### 4. Results

#### 4.1. Demographic Characteristics of the Participants

Table 2 indicates some demographic characteristics of outpatients who participated in the psychometric assessment of the IRA-PISI. The mean age of the 257 participants was 37.48 ± 13.84 (range: 18 - 75) years. Besides, 53% and 68.09% of the participants were females and married, respectively.

<table>
<thead>
<tr>
<th>Variables</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>137 (53.3)</td>
</tr>
<tr>
<td>Male</td>
<td>120 (46.69)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age (37.48 ± 13.84; range: 18 - 75 years)</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-29 (youth)</td>
<td>94 (36.57)</td>
</tr>
<tr>
<td>30-59 (adult)</td>
<td>154 (59.92)</td>
</tr>
<tr>
<td>≥ 60 (elderly)</td>
<td>9 (3.51)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Marital status</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>74 (28.8)</td>
</tr>
<tr>
<td>Married</td>
<td>175 (68.09)</td>
</tr>
<tr>
<td>Widowed and divorced</td>
<td>8 (3.11)</td>
</tr>
</tbody>
</table>

### Table 2. Demographic Characteristics of Participants in the Psychometric Assessment of the Iranian Rapid Assessment for Psychiatric Illness Screening Instrument
4.2. Validity and Reliability Properties

Table 3 showed that the content validity and reliability of the questionnaire. The initial questionnaire consisted of 21 questions. Seven questions were removed during the validation process by expert opinions (CVI was lower than 0.75). According to the expert opinions, finally, a 14-item screening questionnaire was developed. The minimum and maximum scores for the questionnaire were 14 and 70, respectively. Table 3 indicates the content validity and reliability indices of the questionnaire among 12 psychiatric experts who had high work and research experience. The overall relevance of the questionnaire was established (CVI = 0.91) among 12 experts. Moreover, according to the expert opinions, all items had relevancy over 0.83.

The reliability of the tool was determined in a sample of 30 outpatients presenting to family physicians who responded in two steps at a one-week interval. The ICC was higher than 0.89 in all items. The ICC values of the items/questions ranged from 89.3% to 100%. Furthermore, the overall ICC and Cronbach’s alpha were 96% and 83%, respectively (Table 3).

Figure 1 shows the ROC of the developed psychiatric screening tool and the gold standard. Based on the findings, a score of 28 was the best cutoff point where it had the highest sensitivity and specificity. The lowest and highest questionnaire scores after screening 257 outpatients were 14 and 70, respectively.

After determining the best cutoff point, the screening outcome was categorized as problematic and healthy. The results showed that the sensitivity and specificity of the questionnaire compared to the gold standard were 83.62% and 75.17%, respectively (Table 4).

5. Discussion

This study determined the psychometric properties of the IRA-PISI in PHC settings. Due to the absence of a valid and localized rapid assessment tool for screening common PDs in first-line PHC settings, this questionnaire can be a concise, easy-to-administer screening tool to promote the specificity of early detection among first-line healthcare providers, especially family physicians. Furthermore, early detection and treatment of PDs by family physicians and first-line healthcare providers have many challenges due to various types of PDs, poor education and information about mental health among outpatients, face-to-face interviews, and being subjective features feathers of PDs (13, 14).

Diagnoses are sometimes missed given the time demands of clinical practice. Screening tools and scales have been suggested to improve accurate diagnosis. A problem with most study efforts on screening scales is the confusion between diagnostic instruments, tests, and screening. From a clinical perspective, it is most important that a diagnostic tool has high and acceptable sensitivity so that most cases are detected since more time-intensive/expensive follow-up diagnostic inquiries will presumably only occur in patients who are positive on the initial screening. Physicians and healthcare providers vary in how they analyze data in identifying mental disorders. The present instrument could be used for screening, early detection, and minimizing missed cases (15, 16). Family physicians frequently diagnose and treat PDs, particularly in patients enrolled in care plans. One study revealed that 25% to 30% of patients presenting to primary care physicians have PDs (17). Investigations showed that 30 to 80% of primary care patients are undetected by primary care physicians (18). To avoid missing a psychiatric diagnosis, family physicians should use a systematic approach when assessing a patient for PDs, such as a valid tool with high sensitivity (19).

The current instrument aids first-line healthcare staff to provide an appropriate framework and path to screen and recognize common psychiatric disorders such as anxiety and mood disorders in primary care patients in PHC or any medical setting. The initial diagnosis of PDs by general medical doctors and other healthcare providers remains uncertain, and a significant proportion of psychiatric patients are undiagnosed or misdiagnosed.

In Iran, there are various psychometric studies for the validation of mental disorders, including the Templar Death Anxiety Scale-Extended (20), the Persian Version of Type D Personality Scale (DS14) (21), The Young Early Mal-adaptive Schema Questionnaire (YSMQ) (22), among others (23-25). However, the findings revealed that the screening data may be more readily integrated when specific measures are required and low additional computation on part of the clinician is needed.

It is estimated that PHC physicians did not routinely identify patients with PDs since it is needed to specific field. Furthermore, algorithm-based diagnoses are more complex than tools based on a simple sum of scores (26). The findings indicated that PHC providers have a shortage in the identification of PDs due to reasons including physical complaints as the major objective of PHC visits, psychiatric stigma, and time constraints (27). Therefore, rapid screening tools with the simple sum of scores can go a long way toward solving these problems (10).

Content validity must be ensured when developing a tool. On the other hand, content validity is a prerequisite for other validities and has a close relationship with reliability. Therefore, evaluating the content validity is very important in designing questionnaires (28). The process of quantitative validating using the opinions of experts and participants makes questionnaires with high relevance, transparency, and comprehensiveness. In this study, the
Table 3. Comparison of Items, Overall Relevancy, and Reliability of Developed Tool (14 items) for Screening Common Psychiatric Disorders in Primary Health Care Settings

<table>
<thead>
<tr>
<th>No.</th>
<th>Question/Item</th>
<th>Content Validity (Relevancy)</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of Observed</td>
<td>Overall</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agreements Among 12 Experts</td>
<td>Relevancy</td>
</tr>
<tr>
<td>1</td>
<td>Have you had any problems at work recently?</td>
<td>12</td>
<td>100%</td>
</tr>
<tr>
<td>2</td>
<td>Have you had communication problems with friends or colleagues or family members recently?</td>
<td>12</td>
<td>100%</td>
</tr>
<tr>
<td>3</td>
<td>Have you had depressed feelings or poor moods over the past two weeks?</td>
<td>11</td>
<td>91%</td>
</tr>
<tr>
<td>4</td>
<td>Have you felt a decrease in life enjoyment or futility over the past two weeks?</td>
<td>10</td>
<td>83%</td>
</tr>
<tr>
<td>5</td>
<td>Have you felt low energy, sluggish or tired too quickly over the past two weeks?</td>
<td>11</td>
<td>91%</td>
</tr>
<tr>
<td>6</td>
<td>Have you felt anxious and apprehensive about various issues during the last six months?</td>
<td>12</td>
<td>100%</td>
</tr>
<tr>
<td>7</td>
<td>Do you feel you have to repeat some works or things over and over again?</td>
<td>12</td>
<td>100%</td>
</tr>
<tr>
<td>8</td>
<td>Do you have any annoying thoughts that are difficult for you to get rid of?</td>
<td>11</td>
<td>91%</td>
</tr>
<tr>
<td>9</td>
<td>Have you experienced the weirdness or ambiguity of the people or environment around you?</td>
<td>10</td>
<td>83%</td>
</tr>
<tr>
<td>10</td>
<td>Do you have beliefs that seem strange or unacceptable to those around you?</td>
<td>11</td>
<td>91%</td>
</tr>
<tr>
<td>11</td>
<td>Have you wished you had not been born or not been alive over the past few weeks?</td>
<td>12</td>
<td>100%</td>
</tr>
<tr>
<td>12</td>
<td>Have there been times in your life when you felt very energetic or overjoyed?</td>
<td>12</td>
<td>100%</td>
</tr>
<tr>
<td>13</td>
<td>Have there been times in your life when you could not control your anger?</td>
<td>12</td>
<td>100%</td>
</tr>
<tr>
<td>14</td>
<td>Are there times in your life when you feel your body needs sleep much less than ever?</td>
<td>11</td>
<td>91%</td>
</tr>
</tbody>
</table>

Table 4. Sensitivity and Specificity of the Questionnaire Compared to the Gold Standard

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Gold Standard (DSM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Positive</td>
<td>97</td>
</tr>
<tr>
<td>Negative</td>
<td>19</td>
</tr>
</tbody>
</table>

\(^a\) Sensitivity = TP/(TP + FN) = 97/97 + 19 = 83.62%.

\(^b\) Specificity = TN/(TN + FP) = 106/106 + 35 = 75.17%.

The screening tool had acceptable validity and reliability for screening common PDs. The findings revealed that the overall content validity of the tool was 91%, indicating the high agreement between experts and confirming the relevance and appropriateness of the tool. The minimum acceptable value of content validity for the overall tool is 80% (29).

The sensitivity of tools, especially screening tools, is another major criterion of validity. The sensitivity index should be maximized in the first line of the healthcare system (30, 31). The sensitivity and specificity of this questionnaire compared to the gold standard (psychiatric interviews with a structured DSM instrument) were 83.62% and 75.17%, respectively (32).

Regarding the reliability of the questionnaire, internal reliability (Cronbach’s alpha coefficient) and temporal reliability (ICC) were higher than 0.89 in all questions. Consequently, overall ICC and Cronbach’s alpha coefficient were reported as 0.96 and 0.83, respectively. The results showed that all scales and subscales had high reliability. The internal consistency of all scales and subscales in this study was very high. This means that the items in each test or subscale were correlated with each other, as well as with all
items. The highest reliability (ICC) among the questions was related to question 11 (suicide ideation), and the lowest reliability was related to question 9 (psychotic disorders).

There is limited evidence on developing and validating a screening tool in PHC (33). This study is unique as it provides family physicians in PHC settings with a rapid questionnaire. Family physicians frequently identify and treat PDs, particularly in patients enrolled in care programs. It is estimated that 30% to 80% of PDs are undetected by PHC providers (17, 18). Most outpatients with mental disorders are deprived of mental healthcare due to inappropriate case management by first-line healthcare providers, unavailability of valid, reliable, and structured tools, poor awareness of first-line healthcare providers, patient orientation to the private sector, and strong private sector (2, 34). Although most first-line healthcare providers have poor information about PDs, this questionnaire provides a rapid assessment survey for screening common PDs in a short time. On the other hand, completing the questionnaire by the interviewer or patients is possible.

5.1. Limitations

This study is a cross-sectional study that could not assess sensitivity to change over time or responsiveness to changes. Longitudinal studies and high sample size are
required to assess this validity. The use of tools in different ethnicities from various cities requires a pre-test study to ensure the feasibility of using the tool in these populations.

5.2. Conclusions

Findings indicated that the Persian version of IRA-PISI (14 items) showed high validity and sensitivity for screening common PDs in the first-line healthcare system of Iran. Given that the patient load is high for family physicians, they can use this rapid and efficient questionnaire without spending much time. Investigations are needed to evaluate whether such a rapid and valid screening questionnaire is generalizable to other contexts and health systems, and how PDs screening should be best introduced into routine health care practice.

Acknowledgments

This study was extracted from Mostafa Mostofi’s thesis, MD, a student in psychiatry. We express our gratitude to our colleagues in the Research Center of Psychiatry and Behavioral Sciences and the Development Research Unit of Razi Psychiatric Hospital, Tabriz University of Medical Sciences.

Footnotes

Authors’ Contribution: AF and MF contributed to protocol development and interpretation of data and drafted all sections of the manuscript. HA prepared the first draft of the article and conducted data analysis and measurements. MM conceived the study, participated in the design and coordination of measurements, performed the measurements, and assisted in drafting. SS contributed to the data collection, design, and development of the questionnaire.

Conflict of Interests: The authors declare no conflict of interest or financial disclosures.

Data Reproducibility: It was not declared by the authors.

Ethical Approval: This study was extracted from the thesis of Dr. Mohsen Mostofi, a psychiatry student, and approved by the Ethics Committee of Tabriz University of Medical Sciences (IR.TBZMED.REC.1397.186) (ethics.research.ac.ir/EthicsProposalView.php?id=13223).

Funding/Support: This study was funded by the Department of Psychiatry, Tabriz University of Medical Sciences.

Informed Consent: Written informed consent was obtained before the interviews.

References


Iran J Psychiatry Behav Sci. 2022;16(2):ei12790.


