Screening bipolar disorders in a general hospital: Psychometric findings for the Persian version of mood disorder questionnaire and bipolar spectrum diagnostic scale

Amir Shabani1, Moslehi Mirzaei Khoshalani*2, Seyyedreza Mahdavi2, Masoud Ahmadzad-Ast3

Received: 24 Oct 2016 Published: 28 May 2019

Abstract

Background: Mood Disorder Questionnaire (MDQ) and Bipolar Spectrum Diagnostic Scale (BSDS) are used to screen patients with bipolar disorders and have been examined in some psychiatric settings. The present study aimed to assess the validity and reliability of these 2 tools on inpatients in a general hospital.

Methods: In a cross-sectional study in 2011, a total of 207 inpatients admitted to different wards of Rasoul Akram hospital, Tehran, were selected by systematic random sampling. Demographic questionnaire, MDQ, and BSDS were completed. Also, Structured Clinical Interview for DSM-IV axis I disorders (SCID-I) was performed for all participants within 72 hours. The SCID-I was used as the gold standard of psychiatric diagnoses to identify the predictive validity of the 2 screening tools. Sensitivity and specificity indices were identified using Roc curve. The 2 screening tools were recomputed by 20% of the patients (n=43) after 3-7 days to measure test-retest reliability using paired t test and correlation between measures in 2 separate occasions.

Results: In this study, 101 female and 106 male (m=36.9±15.5 yrs.) patients were entered the study, of them 56 (32 males) had bipolar disorder according to SCID-I. The most common bipolar disorder was bipolar disorder type II (9.7%). Pearson’s test showed a high test-retest reliability for both MDQ (r=0.72, p<0.001) and BSDS (r=0.77, p<0.001). For MDQ, the scores 5 (sensitivity=0.60; specificity=0.73) and 6 (sensitivity=0.56; specificity=0.77) were the best cutoff points. Positive and negative predictive values for the mentioned cutoff points were 0.45 and 0.83 (for the score 5) and 0.48 and 0.82 (for the score 6), respectively. The best cutoff point for BSDS was 11 with the sensitivity, specificity, and positive and negative predictive values of 0.74, 0.69, 0.47, and 0.87.

Conclusion: The Persian versions of MDQ and BSDS have acceptable validity and reliability to screen Persian patients with bipolar spectrum disorders in a general hospital.

Keywords: Bipolar disorder, Screening, Mental disorders, Bipolar spectrum diagnostic scale, Mood disorder questionnaire, Sensitivity and specificity, Psychometrics

Corresponding author: Moslehi Mirzaei Khoshalani, mirzaei.m@iiums.ac.ir

Introduction

Bipolar disorder (BD) is a recurrent, chronic, disabling, and costly psychiatric disorder and is considered the eighth leading cause of medical disability worldwide (1, 2). The 12-month and lifetime prevalence of standard BD was estimated about 1.8% and 4.1%, respectively, (3) and of subthreshold of about 6.4% (4) in the community.

†What is “already known” in this topic:
Mood Disorder Questionnaire (MDQ) and Bipolar Spectrum Diagnostic Scale (BSDS) are reliable and valid tools in some settings.

—What this article adds:
This study demonstrated acceptable validity and reliability of Mood Disorder Questionnaire (MDQ) and Bipolar Spectrum Diagnostic Scale (BSDS) to screen inpatients with bipolar disorders in a general hospital.
Patients with bipolar disorder have high rates of general medical conditions. According to a cross national epidemiological investigation in a community sample, people with bipolar disorder (the group with a positive manic episode screen) had a prevalence increment for migraine, asthma, chronic bronchitis, gastric ulcer, multiple chemical sensitivities, and chronic fatigue syndrome. Comorbidity with these medical conditions is associated with poor psychosocial functioning and may result in increased premature mortality possibly due to detrimental lifestyle, biological indices, side-effects of medications, and unevenness in utilization of health care (6). Although bipolar disorder patients experience relative under recognition of and inattention to their many physical diseases in mental health settings (7), the prevalence of mood disorders in general hospital settings is quite significant [17% based on Elisei et al study (8)]; and according to National Hospital Discharge Survey (NHDS), there is an average 10% per year rise in the proportion of discharges with BD (9).

The considerable high prevalence of general medical conditions in BD can result in higher mortality and health problems, and despite inequalities in higher health service use, it can be speculated that screening BD in patients referring to general hospitals due to medical conditions could be a backward screening method for BD and may result in more effective interventions by early detection of BD in those patients. Developing and implementing an easy to use and valid tool for this purpose would be a useful intervention in health of BD patients.

Until recently, there were no validated research tools for BD in Persian language (Farsi). Fortunately, validating studies in recent years have presented psychometric values of some useful instruments on Iranian patients, including Young Mania Rating Scale (10), Hypomania Check-List 32 (11), Bipolar Depression Rating Scale (12), Mood Disorder Insight Scale (13), Mood Disorder Questionnaire (MDQ) (14), and Bipolar Spectrum Diagnostic Scale (BSDS) (14). The 2 latter ones are self-report questionnaires used to screen patients with BD.

The MDQ has been widely used in different languages, various settings, and different populations (15). Review of Zimmerman and Galione (16) showed that MDQ has high sensitivity and positive predictive value in screening BD among psychiatric outpatients, and low specificity and negative predictive value in general population. Also, its sensitivity to detect bipolar disorder type I (BID) is higher than bipolar disorder type II (BIID) (16). The original study (17) on MDQ demonstrated the cutoff score of 7 in detecting BD, but later studies led to inconsistent findings (16).

After the MDQ, another screening tool – the BSDS – was introduced to cover milder forms of BD (18). BSDS has been studied less than MDQ, but it seems to be equally sensitive to find BID, BIID, and bipolar disorder- not otherwise specified (19); the point that may compensate for the drawbacks of MDQ.

Several factors that are naturally different in various settings can affect the performance of bipolarity screening tools. Prior probabilities, BD subtype, and the sample source are among these factors (16, 20). Thus, it is required to identify the psychometric properties of the screening tools in the target setting before applying them. Due to the high prevalence (21, 22) and frequent underdiagnosis (22) of mood disorders among patients admitted to general hospitals, the present study was conducted to clarify the psychometric values of the MDQ and BSDS in a new target population.

**Methods**

**Participants and sampling**

This cross sectional study was conducted on 207 patients hospitalized in different wards of Rasoul Akram general hospital, affiliated to Iran University of Medical Sciences in Tehran, Iran, in 2011. The patients (n=207) were selected from those who were admitted to 11 wards of Internal Medicine, General Surgery, Orthopedics, Dermatology, Ophthalmology, Nephrology, Neurology, Psychiatry, Pain, Neurosurgery, and Cardiology.

The sample volume in each ward was identified according to the ratio of the admitted patients, and systematic random sampling was used to select patients in each ward, i.e., 1 patient was selected among every 4 sequential patients. Also, gender matching was done among the samples of different wards.

The inclusion criteria were as follow: age 18-65 years, speaking Farsi, and providing informed consent. The patients with low alertness and poor physical conditions in the interview and completing the questionnaires were excluded.

**Instruments**

**The Structured Clinical Interview for DSM-IV axis-I disorders (SCID-I):** The SCID-I is the most widely utilized semi-structured diagnostic interview for identifying the disorders labeled as the axis-I of the DSM-IV (23). In the present study, the clinician version of the Persian SCID-I (SCID-CV) was used. The normalization study of the Persian SCID-CV has demonstrated fair to good test-retest reliability for most diagnostic classes (24, 25).

**The Bipolar Spectrum Diagnostic Scale (BSDS):** The BSDS is a self-assessment tool developed and revised by Pies, Ghaemi, and Miller (18). In the first part of this scale, there are 19 sentences which narrate the main symptoms of bipolar spectrum disorders. At the second part, there is a discriminant item that assesses how well the first part fits to their experiences. Scores on the second part can vary from 0 (doesn’t describe me at all) to 6 (fits me perfectly). So, the score of BSDS varies from 0 to 25. According to the original study, the sensitivity and specificity of BSDS for screening bipolar disorders was 0.76 and 0.85, respectively. BSDS has been validated in a number of countries, including Korea (26), Iran (14), Spain (27), China (28), and Argentina (29).

The first validation study of BSDS in 181 Iranian outpatients by Shabani et al (14) showed that at the score of 14, the sensitivity, specificity, and positive and negative predictive values of the Persian BSDS were 0.52, 0.79, 0.81, and 0.49, respectively.

**The Mood Disorder Questionnaire (MDQ):** The MDQ
is a brief self-assessment tool which screens the lifetime history of bipolar disorder. The questionnaire includes 3 sections. The first section with 13 yes/no questions screens manic/hypomanic symptoms of DSM-IV. A yes/no question in the second section aims at clarifying the clustering of symptoms and the final question measures the level of impairment in functioning from “no problem” to “serious problem”.

In the original validation study of MDQ (17), which was administered on 198 mood disorder outpatients, a screening score of 7 and more yielded an acceptable sensitivity (0.73) and a high specificity (0.90).

MDQ has been validated in different countries. After its development in the United States, MDQ has been translated and used in more than 10 countries [China (30), Hong Kong (31), Argentina (29), Brazil (32), Finland (33), France (34), Italia (35), Korea (36), UK (37), Spain (38), and Turkey (39)] including Iran (14) and still continues in other countries (40). Some studies have applied it in different target populations, and mostly on mood disordered psychiatric out- and inpatients. Diverse settings have also been studied either for screening or case finding aims. To draw a conclusion about the MDQ’s performance, Zimmerman’s critical review, which showed that the MDQ had modest sensitivity and good specificity across all studies, should be considered. MDQ is more sensitive in patients than in the general population and in bipolar I disorder than bipolar II disorder (16).

Procedure

The demographic data were first filled out and then MDQ and BSDS were administered on the probands. The items of the 2 self-report scales were read loudly by a researcher and the patients’ responses were registered. In a time interval not more than 72 hours, a clinical interview (SCID-I) was performed by a trained resident of psychiatry. The training was provided by the first author.

To assess the test-retest reliability of the MDQ and BSDS, they both were used again in 3 to 7 days after the first time through a telephone contact on 43 probands.

Statistical analysis

Concurrent validity of MDQ and BSDS against each other was calculated. The SCID-I interview was used as the gold standard of psychiatric diagnoses to identify the predictive validity of the 2 screening tests. Sensitivity and specificity indices were identified using Roc curve.

The Pearson’s correlation test was used to examine the correlation between the MDQ and BSDS. P value<0.05 was considered significant. Test-retest reliability was estimated by paired t test and correlation between measures in 2 separate occasions. Statistical analyses were performed using the SPSS for Windows software version 16, and statistical significance was considered while having a 2-tailed p<0.05.

Results

The sample consisted of 101 female and 106 male patients hospitalized in all wards of Rasoul Akram general hospital, except for the Emergency and Pediatric wards. The mean age (±SE) was 36.9 (±15.51) years. A total of 18.8% had elementary school education level. Other levels were 29.9% (middle school), 30% (high school), 5.3% (diploma), and 10.6% (university degrees). In terms of employment status, 54.1%, 7.2%, 32.9%, and 5.8% of the participants were employed, unemployed, housewives, and retired, respectively.

A total of 56 patients out of 207 (27.1%) were diagnosed with bipolar disorder according to the SCID-I (25 females and 32 males): bipolar disorder type II 9.7% (n=20), bipolar disorder type I 8.7% (n=18), bipolar disorder-NOS 4.8% (n=10), and cyclothymic disorder 3.9% (n=8).

Pearson test results showed high rates of test-rest reliability for both MDQ (paired correlation R=0.72, p<0.001, and paired t-test, p=0.882; n=43) and BSDS (R=0.77, p<0.001, and paired t test, p=0.078, n=43). Considering internal consistency, the Cronbach’s alpha coefficients were 0.745 and 0.861 for the MDQ and BSDS, respectively. The correlations between each item of the MDQ and BSDS with its total score were significant (p<0.01), ranging from r=0.27 (the item 11 of the question 1) to r=0.68 (the question 3) for the MDQ, and r=0.18 (the item 18) to r=0.90 (the item 20) for the BSDS.

The Sensitivity and specificity of MDQ and BSDS to screen all types of bipolar disorders are shown in Figures 1

Fig. 1. Sensitivity and specificity of the MDQ to screen all cases of bipolar disorder

http://mjiri.iums.ac.ir
Med J Islam Repub Iran. 2019 (28 May); 33.48.
Screening bipolar disorders among patients in a general hospital

and 2. The sensitivity and specificity of MDQ at the score 5 were 0.60 and 0.73 and at the score 6 were 0.56 and 0.77, respectively. These scores were the best cutoff points among other scores (Table 1). After using these thresholds, the positive and negative predictive values of MDQ were 0.45 and 0.83 (for the score 5), and 0.48 and 0.82 (for the score 6), respectively. According to Table 2, the best cutoff points for screening bipolar disorders in all types are the scores 10, 11, and 12, depending on the situation in which a user wants to have a screening instrument with good sensitivity or enough specificity. The positive and negative predictive values of the BDSs at the above cutoffs were 0.44 and 0.91 (for the score 10), 0.47 and 0.87 (for the score 11), 0.49 and 0.86 (for the score 12), respectively (Fig. 2).

Discussion
To the best of our knowledge, this was the first psychometric study on MDQ and BDSs with general hospital inpatients as the target. MDQ had a sensitivity of 0.60 and specificity of 0.73 at the best cutoff point, compared with SCID-I as a gold standard diagnosis tool. This sensitivity is quite more than what was found in the general population (0.281) (42) and slightly less than hospitalized patients with comorbidity (0.84) (43) and almost equal to MDD outpatients in family medicine clinic (0.58) (44). This finding can be explained by the fact that in general hospitals or primary care settings it is expected to find more signs and symptoms of mood disorders than in community or general population settings; on the other hand, it can be reasonably expected to see more frequent and severe features of mood disorders in psychiatric populations (20).

Early detection of BD is of great importance for some reasons. First, delay in diagnosing BD results in poorer prognosis (45) with higher rates of suicidal attempt, suicide, recurrence, chronicity, and treatment resistance (46, 45). Second, considering difficulty in detecting hypomania (47), poor insight of the patients (48, 49), much higher frequency of depressive episodes than manic/hypomanic episodes (50, 51), and diagnosis of BD could be very difficult. Therefore, some screening instruments have been developed to make early detection of BD easier (20).

Some studies have compared screening tools that detect BD; for example, Sasdelli et al (52) have shown that MDQ is more accurate and feasible in recognizing BD in patients with depressive symptoms. Also, there is ample evidence of underdiagnoses and misdiagnosis of BD that may lead to both missing some subjects in different medical settings such as general hospitals and to undertreatment of BD comorbid conditions. These comorbidities may result in additional costs and health consequences. As practitioners in general medical settings are not completely familiar with classic BD diagnostic methods, and considering that in those settings the main complaints of the subjects are far from mood problems such as BD, it could be assumed that introducing and validating accurate, brief, and feasible tools of BD screening can have substantial clinical value for both practitioners and patients.

BDSs has good sensitivity (0.74) and specificity (0.69), but like MDQ it has a low to moderate positive predictive value (0.47) and high negative predictive value (0.87). The original BDS study by Ghaemi et al showed that the sensitivity, specificity, and NPV were more than 0.75. PPV was around 0.36 comparable to our study (18). A notable similarity was low PPV in both studies which in-

**Table 1. Sensitivity and specificity of the MDQ for each final score**

<table>
<thead>
<tr>
<th>Score</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>60</td>
<td>66</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>69</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>73</td>
</tr>
<tr>
<td>6</td>
<td>56</td>
<td>77</td>
</tr>
<tr>
<td>7</td>
<td>51</td>
<td>85</td>
</tr>
<tr>
<td>8</td>
<td>42</td>
<td>91</td>
</tr>
<tr>
<td>9</td>
<td>30</td>
<td>94</td>
</tr>
</tbody>
</table>

**Table 2. Sensitivity and specificity of the BDSs for each final score**

<table>
<thead>
<tr>
<th>Score</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>88</td>
<td>43</td>
</tr>
<tr>
<td>9</td>
<td>86</td>
<td>49</td>
</tr>
<tr>
<td>10</td>
<td>84</td>
<td>59</td>
</tr>
<tr>
<td>11</td>
<td>74</td>
<td>69</td>
</tr>
<tr>
<td>12</td>
<td>68</td>
<td>73</td>
</tr>
<tr>
<td>13</td>
<td>60</td>
<td>77</td>
</tr>
<tr>
<td>14</td>
<td>53</td>
<td>79</td>
</tr>
<tr>
<td>15</td>
<td>49</td>
<td>85</td>
</tr>
</tbody>
</table>
dicated the risk of false positive, a point that has been demonstrated and discussed in Ghaemi's review over bipolar self-report screening tools (20). These characteristics may reveal the potential of aforementioned tools as screening tools where further diagnostic procedure is needed to confirm diagnosis, and using them as diagnostic tools may find high rates of false positives in patients with BD diagnosis.

All studies that have validated Chinese, Spanish, Persian, and Argentinian versions of BSDS have been performed on psychiatric outpatients. Using different optimal cutoff points, ranging from 11 (19) to 14 (14), resulted in 0.52 (Persian) to 0.74 (Chinese) sensitivity and 0.79 (Persian) to 0.97 (Chinese) specificity. However, Zimmerman et al through the evaluation of BSDS performance on a representative sample (1100 outpatients) showed that the scale has a very good NPV (98.0%), but very low PPV (0.16) to be a suitable screening tool (19). This characteristic has made this test vulnerable to overdiagnosing BD.

Conclusion

This study demonstrated that MDQ and BSDS have acceptable psychometric properties as screening measures for patients with bipolar spectrum disorders admitted in general hospitals. However, some limitations should be addressed. The psychiatric diagnoses were done by a resident of psychiatry. Although she had been trained to perform SCID, overdiagnosis might have happened. Also, the potential impact of hospitalization reason and disorder type on tests characteristics were not assessed in this study that may influence responses on some assessment domains, and they should be considered in future studies. Furthermore, in this study, the participants were recruited from patients admitted to a university governed hospital, and thus the findings may not be easily generalizable to other patients in private hospitals. However, this was the first study to examine the performance of MDQ and BSDS in nonpsychiatric inpatients and needs to be replicated. Moreover, the PPV of these tools need to be improved.

Acknowledgments

This study was supported by a grant from the Mental Health Research Center, Iran University of Medical Sciences, Tehran, Iran. The authors appreciate the kind cooperation of Dr. Mehdi Nasr Esfahani in Rasoul Akram hospital. Also, all the probands who kindly took part in the program are appreciated. Sincere thanks go to Dr. Sona Setayesh and Zahra Farahmand for conducting research questionnaires.

Conflict of Interests

The authors declare that they have no competing interests.

References

25. Shoshtari MH, Davari-Ashtiani R, Shahrivar Z, Shabani A, Sem-

A. Shabani, et al.

http://mjiri.iums.ac.ir