Development and validation of a screening instrument for bipolar spectrum disorder: The Mood Disorder Questionnaire Thai version

Punjaporn Waleeprakhon1
Pichai Ittasakul1
Manote Lotrakul1
Pattarabhorn Wisajun1
Sudawan Jullagate1
Terence A Ketter2

1Department of Psychiatry, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand; 2Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA

Background: The Mood Disorder Questionnaire (MDQ) has been translated to many languages and has been used in many countries as a screening instrument for bipolar disorder. The main objective of this study was to evaluate validity of the Thai version of the MDQ as a screening instrument for bipolar disorder in a psychiatric outpatient sample, and to determine its optimum question #1 item threshold value for bipolar disorder.

Methods: The English language Mood Disorder Questionnaire (MDQ) was translated into Thai. The process involved back-translation, cross-cultural adaptation, field testing of the prefinal version, as well as final adjustments. Two hundred and fifty major depressive disorder outpatients were further assessed by the Thai version of the MDQ and the Thai version of the Mini International Neuropsychiatric Interview (MINI). During the assessment, reliability and validity analyses, and receiver operating characteristic curve (ROC) analysis were performed.

Results: The Thai version of the MDQ screening had adequate internal consistency (Cronbach’s alpha = 0.791, omega total = 0.68, and omega hierarchical = 0.69). The optimal question #1 item threshold value was at least five positive items, which yielded adequate sensitivity (76.5%), specificity (72.7%), positive predictive value (74.3%), and negative predictive value (75.0%). The ROC area under the curve (AUC) for this study was 0.82 (95% confidence interval: 0.70 to 0.90).

Conclusion: The Thai version of the MDQ had some useful psychometric properties for screening for bipolar disorder in a mood disorder clinic setting, with a recommended question #1 item threshold value of at least five positive items.

Keywords: Thai MDQ, screening test, reliability, predictive validity, cutoff point

Introduction

Bipolar disorder is a common, serious, recurrent illness marked by episodes of depression and mood elevation, and entails serious psychosocial consequences (eg, suicide, incarceration, bankruptcy, divorce, employment termination, occupational disability, and diminished social function).1 It is one of the leading causes of worldwide disability, especially in those aged 15–44 years.2

Accurate diagnosis of individuals with bipolar disorder can be challenging because people with bipolar disorder commonly present in the more pervasive depressive phase and may not recall previous manic or mixed episodes, or may not be able to distinguish them from prior major depressive episodes, raising the risk that they may receive an inaccurate diagnosis of unipolar major depressive disorder.3 For example, in one study, approximately 70% of patients with bipolar disorder were misdiagnosed,4 with the most common incorrect initial diagnosis being unipolar
major depressive disorder. Misdiagnosis can lead to inappropriate treatment and poor treatment outcome. For example, standard antidepressants, which commonly provide adequate efficacy and tolerability in unipolar major depressive disorder, when administered to individuals with bipolar disorder, can have inadequate efficacy (leaving individuals depressed) and/or tolerability (causing emergence of manic symptoms). Indeed, investigators have reported that as many as over one-third of bipolar disorder patients may endure as long as a decade of affective symptoms before appropriate diagnosis and treatment.

Hirschfeld et al developed the Mood Disorder Questionnaire (MDQ), a brief self-report screening instrument for bipolar spectrum disorders; the MDQ detects past episodes of mania or hypomania via 13 yes/no items derived from both the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria and clinical experience. In clinical settings, the MDQ has had good sensitivity and specificity (73% and 90% respectively), and it has been translated to multiple languages and validated in multiple countries. Although the sensitivity of the MDQ may be limited in community settings, multiple studies have showed that the MDQ is a valid clinical screening instrument for bipolar disorder, even in international settings.

However, the MDQ has not been translated into and validated in the Thai language. In the current study, the objective was to evaluate the validity of the Thai version of the MDQ as a screening instrument for bipolar disorder in major depressive disorder outpatients, and to determine its optimum question #1 item threshold value for bipolar disorder.

Methods

Study design

The study was conducted at an outpatient psychiatric clinic of Ramathibodi Hospital, Bangkok that primarily treats general psychiatric patients. The protocol was approved by the Ethics Committee on Human Experimentation of the Faculty of Medicine, Ramathibodi Hospital, Bangkok, Thailand. All subjects provided verbal and written informed consent prior to participation.

Participants

The participants were native Thai-speaking adults (age ≥18 years) recruited between October 1, 2012 and January 31, 2014 from the psychiatric outpatient clinic at Ramathibodi Hospital; all had a clinical diagnosis of unipolar major depressive disorder, determined by a psychiatrist using the DSM IV, Text Revision (DSM-IV-TR) criteria. Participants with a psychiatric or physical disorder that prevented them from being interviewed or undermined their ability to provide accurate information, and those who declined participation in the study or refused to provide informed consent, were excluded.

Measures

The MDQ is a self-report measure for a lifetime history of mania or hypomania, consisting of 13 yes/no symptom questions based on the DSM-IV criteria for bipolar disorder. The symptom questions are followed by a single yes/no question about whether symptoms clustered during the same period of time. The final question evaluates the level of impairment resulting from the symptoms, with rating on a four-point scale (no problem, minor problem, moderate problem, or serious problem). After obtaining permission from the copyright holder, the MDQ was translated into Thai from the original English MDQ. The MDQ was translated into Thai according to the guidelines for cross-cultural adaptation of self-report measures. The process included two independent forward translations of the original English MDQ into Thai, consensus between translators on the forward translation, back-translation by a bilingual English teacher, and a review of the back-translation by the authors. Ten patients attending the outpatient department were invited to complete and to give comments on the prefinal version. Final modifications and adjustments were made accordingly.

The Mini International Neuropsychiatric Interview (MINI), Version 5, is a standardized clinical diagnostic interview schedule for DSM-IV Axis-I disorders. It can be reliably administered by lay interviewers who have appropriate training. The Thai version of MINI (which was translated from the English version of the MINI, Version 5) was used in this study as the “gold standard” diagnostic tool for identifying the presence of bipolar disorder.

Procedure

Subjects were recruited at their scheduled clinic visit, during which they completed the Thai version of the MDQ. Demographic data (eg, sex, date of birth, religion, marital status, education, and occupation) were recorded. After completing the questionnaires, participants were then assessed for bipolar disorder by two research assistants who were unaware of the participants’ MDQ results. The research assistants were trained to use the Thai version of MINI. Interrater reliability of scoring was assessed. The agreement between raters was satisfactory (Kappa = 0.91).
Statistical analysis
All statistical analysis was performed by using IBM SPSS Statistics for Windows, Version 21.0 (IBM Corp., Armonk, NY, USA) and R and the Psy Package Version 3.1.0 (Northwestern University, Chicago, IL, USA). Descriptive and analytic statistics were compiled. Cronbach’s alpha coefficient and McDonald’s omega were used to assess the internal consistency of the scale. The receiver operating characteristic (ROC) curve was plotted to assess the screening performance of the questionnaire. Its accuracy was calculated in terms of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and likelihood ratio for each possible cutoff, and the method of linear interpolation was used to calculate the sensitivity and specificity for each actually possible cutoff (number of positive answers). The optimal question #1 item threshold value was determined by maximizing the Youden’s index. For all analyses, a significant threshold of \( P<0.05 \) was used.

Results
Sample description
A total of 250 outpatients with a clinical diagnosis of unipolar major depressive disorder completed the Thai version of the MDQ and were assessed using the Thai MINI. The mean age (standard deviation [SD]) of participants was 46.9 (13.8) years, and 79.2% of participants were female. Demographic data are provided in Table 1.

Reliability and item analysis
The Cronbach’s alpha coefficient for the Thai version of the MDQ was 0.791. The McDonald’s omega total and omega hierarchical were 0.68 and 0.69 respectively. The frequency of endorsement of MDQ items ranged from 4.8% to 50.8%. The items that were endorsed most frequently were being “easily distracted” (50.8%), having “decreased need for sleep” (41.2%), and having “irritable mood” (36.8%). The items that were endorsed the least were having “increased interest in sex” (4.8%) and being “more social and outgoing” (6.4%). The corrected item-total correlations ranged from 0.24 to 0.57. All items, if deleted, would consistently decrease the total scale alpha (Table 2).

Validity analysis
In 250 participants who were diagnosed with major depressive disorder, a MINI diagnosis of bipolar disorder was given to 60 participants (24.0%) (19 [7.6%] bipolar I disorder, and 41 [16.4%] bipolar II disorder).

As applied in the validity study of the original (English) version of the MDQ, when the response to the second question was positive and the third question affirmed moderate to severe severity, the item threshold values for the first question, which includes 13 subitems, were determined by ROC analysis. Area under the curve (AUC) was 0.82 (95% confidence interval [CI]: 0.70–0.90, \( P<0.0001 \)) (Figure 1).

The sensitivity, specificity, PPV, NPV, and likelihood ratio for the Thai version of the MDQ using different question #1 positive item thresholds are provided in Table 3. For the Thai version of the MDQ, the sensitivity and specificity at the question #1 item threshold value of at least four positive items were 85.3%, and 51.5%, respectively. Using a question #1 item threshold value of at least five positive items, sensitivity and specificity were 76.5% and 72.7%, respectively. Using a question #1 item threshold value of at least six positive items, sensitivity and specificity were 58.8% and 87.9%, respectively.

By maximizing the Youden’s index, a score of five or more positive items was chosen as the optimal question #1 item threshold value for bipolar disorder as it provided a good balance, with not only adequate sensitivity but also,
adequate specificity. By using this five or more question #1 item threshold, more than seven out of ten people with a bipolar disorder would be expected to be correctly identified by the Thai version of the MDQ, whereas more than seven out of ten of those who did not have a bipolar disorder would be expected to be successfully screened out.

### Discussion

The internal consistency of the Thai version of the MDQ in this study (Cronbach’s alpha coefficient \(=0.79\)) was marginally lower than in the studies of the original English MDQ from the United States (alpha coefficient \(=0.84–0.90\)).\(^{10,17}\) However, the reliability of the Thai MDQ was within the acceptable range as Cronbach’s alpha coefficient was greater than 0.70.\(^2\) The participants in this study had high rates of endorsement of distractibility, decreased need for sleep, and irritable mood, which were consistent with results from previous studies.\(^10,17\)

When the Thai version of the MDQ was examined as a continuous measure, its validity was supported by an AUC value of 0.82 (95% CI: 0.70–0.90), which being greater than 0.80, suggested adequate accuracy. The sensitivity and specificity using a question #1 item threshold value of at least five positive items were also adequate,
Our study has several noteworthy limitations. First, participants were patients already being treated in a university hospital; therefore, the participants may not be representative of the patients usually treated in non-university community settings and were certainly not representative of heterogeneous groups of individuals with and without depression in the community, in general. Second, although the Thai version of the MINI performed well in a validity study,\textsuperscript{19} it is still possible that it overestimated or underestimated the rate of manic/hypomanic symptoms in our study. Third, our somewhat liberal question \#1 item threshold value of at least five versus seven (used in some countries) positive items could have overestimated the prevalence of clinically significant manic/hypomanic symptoms. Fourth, the test-retest reliability of the Thai version of the MDQ was not assessed. Generally, this type of reliability is used for measuring the stability of a scale over time, and it is usually assessed after a short period of time. Fifth, there were substantial limitations on the MDQ’s ability to accurately detect “bipolarity”, including the risks of misdetection of borderline personality disorder\textsuperscript{27} and attention-deficit/hyperactivity disorder\textsuperscript{28} as bipolar disorder. Finally, the ability to detect bipolarity in this study was only assessed in patients who were already professionally diagnosed as having had major depressive episodes; in this regard, our sample had a larger percentage of female subjects and a later onset age compared with most cohorts of patients with bipolar disorder.

**Conclusion**

In summary, the Thai version of the MDQ had some useful psychometric properties for screening for bipolar disorder in a mood disorder clinic setting, with a recommended question one item threshold value of at least five positive items. Thus, the Thai MDQ may be a useful instrument to identify a risk for bipolar disorder in Thai clinical settings. However, if the patient screens positive for possible bipolar disorder with the MDQ, the physician should proceed with full clinical evaluation for bipolar disorder.\textsuperscript{29} Further population-based research is needed to assess whether the Thai version of the MDQ would be useful in other settings, particularly as the original English MDQ had limited sensitivity in a (nonclinical) community setting.

**Acknowledgment**

This study was supported by grants from the Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

**Author contributions**

All the Thai authors contributed to the design of the study and wrote the protocol. Pichai Ittasakul and Terence A Ketter managed the literature searches and analyses. Pattarabhorn Wisajun and Sudawan Jullagate were responsible for data collection. Pichai Ittasakul and Punjaporn Waleeprakhon wrote the draft manuscript. Terence A Ketter revised the draft manuscript. Thus, all authors contributed to development of the manuscript, revised it critically for important intellectual content, and gave their approval of this version to be published.

**Disclosure**

The authors report no conflicts of interest in this work.

---

**Table 3** Sensitivity, specificity, PPV, NPV, and LR of the Thai version of the MDQ for diagnosis of bipolar disorder

<table>
<thead>
<tr>
<th>Cutoff points</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Positive LR</th>
<th>Negative LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;3</td>
<td>94.1</td>
<td>45.5</td>
<td>64.0</td>
<td>88.2</td>
<td>1.7</td>
<td>0.1</td>
</tr>
<tr>
<td>&gt;4</td>
<td>85.3</td>
<td>51.5</td>
<td>64.4</td>
<td>77.3</td>
<td>1.8</td>
<td>0.3</td>
</tr>
<tr>
<td>&gt;5</td>
<td>76.5</td>
<td>72.7</td>
<td>74.3</td>
<td>75.0</td>
<td>2.8</td>
<td>0.3</td>
</tr>
<tr>
<td>&gt;6</td>
<td>58.8</td>
<td>87.9</td>
<td>83.3</td>
<td>67.4</td>
<td>4.9</td>
<td>0.5</td>
</tr>
<tr>
<td>&gt;7</td>
<td>47.1</td>
<td>90.9</td>
<td>84.2</td>
<td>62.5</td>
<td>5.2</td>
<td>0.6</td>
</tr>
<tr>
<td>&gt;8</td>
<td>32.4</td>
<td>93.9</td>
<td>84.6</td>
<td>57.4</td>
<td>5.3</td>
<td>0.7</td>
</tr>
<tr>
<td>&gt;9</td>
<td>17.7</td>
<td>97.0</td>
<td>85.7</td>
<td>53.3</td>
<td>5.8</td>
<td>0.9</td>
</tr>
</tbody>
</table>

*Abbreviations:* LR, likelihood ratio; MDQ, Mood Disorder Questionnaire; NPV, negative predictive value; PPV, positive predictive value.