ORIGINAL ARTICLE

Psychiatry and Clinical Psychopharmacology 2021;31(2):173-180
DOI:10.5152/pcp.2021.21425

The Study of Diagnostic Value of Bipolarity Index for Bipolar Disorder in China: Meta-analysis of Sensitivity and Specificity

Sun Fengli[1], Zhu Jianfeng[2], Tao Hejian[3], Jin Weidong[1,3]

[1]Department of Psychiatry, Zhejiang Province Mental Health Center, Zhejiang Province Tongde Hospital, Hangzhou, China; [2]Department of Psychiatry, Second Hospital of Yuhang, Yuhang, China; [3]Zhejiang Chinese Medicine University, Hangzhou, China

ABSTRACT

Objective: Bipolarity index (BI) is one of the diagnostic scales that assist the diagnosis of bipolar disorder (BD), and should be analyzed comprehensively for use in China.

Methods: We searched the Chinese Biomedical Database (CBM), China National Knowledge Infrastructure (CNKI), WANFANG, and Chinese Social Sciences Citation Index (CSSCI) in Chinese to find literature from July 31, 2004 to July 31, 2020, for results related to BI in the diagnosis for bipolar disorder (BD), among which results such as comments, letters, reviews, and case reports were excluded. The rates of sensitivity, specificity, accuracy, positive predictive value, and negative predictive value in diagnosis were synthesized and discussed.

A total of 1237 patients were included in 5 studies. The criteria used for their selection were an anlysis of their results on the BI, and the diagnostic indexes of BI for BD in China.

The sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of BI for BD in China were summarized in every study.

Results: A total of 1237 subjects were included in 5 studies. The random effect model was used to account for the data with RevMan 5.2. The results showed that the diagnostic sensitivity of BI was 0.93 (95% CI: 0.93-1.00), and the specificity was 85% (95% CI: 0.69-0.96). The positive predictive value (PPV) was 74% (95% CI: 0.53-0.91). The negative predictive value (NPV) was 95% (95% CI: 0.81-1.00), and accuracy was 86% (95% CI: 0.77-0.93). Significant heterogeneity was detected across studies regarding these incidence estimates.

Conclusion: The ideal diagnostic value of BI was found, although the studies showed significant heterogeneity. The results must be cautiously and attentively interpreted, in comparison to other diagnostic scales, to perfect the use of BI in clinical psychiatry.

ARTICLE HISTORY

Received: 19 February, 2021
Accepted: 14 April, 2021

KEYWORDS: Bipolarity index, bipolar disorder, diagnostic, meta-analysis

INTRODUCTION

Although the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders) has some significant changes, the sections on missed diagnosis and misdiagnosis of bipolar disorder (BD) remain more or less the same.[1,2] There are 2 patterns and 5 themes. The first model comprises undiagnosed BD, in patients who commonly tend to experience more impulsive emotions and behaviors, thus suffering from the challenges of life. The second model, even considering the possible diagnosis of BD, is diagnosed as other conditions that are similar, such as depression or schizophrenia.[1] Moreover, the first episode of BD is often a depressive episode, which always leads to the diagnosis of major depressive episode (MDD). The diagnosis of BD often lags behind the initial diagnosis of depression, resulting in the delay of accurate diagnosis and treatment of BD. The results of this study show that the early diagnosis of depression is related to a longer delay in diagnosing BD.[2] Therefore, it is important to combine or increase the use of diagnostic scales such as the hypomania checklist-32 (HCL-32), or the mood disorder questionnaire (MDQ) as auxiliary diagnostic tools in order to change or improve the accuracy of diagnosis of BD .[3,4] Both the HCL-32 and the MDQ have higher sensitivity and higher specificity in screening for BD, although there are some slight differences in diagnostic value. At the same time, the bipolar index (BI) scale is also one of the auxiliary diagnostic methods.[5,6] Its application may be more conducive to the diagnosis of BD.[7,8]

In contrast to the HCL-32 and MDQ, the BI scale not only focuses on clinical symptoms or manifestations, but also involves family history, course of illness, characteristics,

Corresponding author: Jin Weidong, e-mail: wdjin@163.com

Cite this article as: Sun FL, Zhu JF, Tao HJ, Jin WD. The study of diagnostic value of bipolarity index for bipolar disorder in china: Meta-analysis of sensitivity and specificity. *Psychiatr Clin Psychopharmacol*. 2021; 31(2): 173-180.

and treatment response.[9,10] This is similar to soft bipolar diagnostic criteria that involve clinical manifestation, age at first onset, family history, personality characteristics, and treatment response.[11,12] The diagnostic criteria for soft BD had been established in China and also been revised after DSM-5 in relation to the bipolarity index (BI).[13] However, the BI seems to have more diagnostic specificity than HCl-32,[14] mainly because it collects more relevant information than the HCL-32 and MDQ which is more closely related to the diagnosis of BD. The results showed that BI has 100% specificity in the diagnosis of BD, which indicates that the BI scale may have more specific significance in the diagnosis of BD.[14] The study findings indicate good reliability and validity for the Chinese version of the BI, which encourages its use as a measure of diagnostic confidence for bipolar spectrum disorders. Further prospective study is necessary to determine whether the BI is useful in identifying subgroups among MDD subjects at high risk for conversion to BP.[6,14]

Chinese researchers have conducted specific studies on different items of the BI scale, and also put forward its diagnostic value.[10,15] They found that the Chinese version of the BI scale has high reliability and validity. With a defined cut-off score, it can help early recognition of BD and provide a new evaluation tool for clinical psychiatry.[15] Since Chinese scholars translated and introduced the BI into China, many clinical studies have been carried out to evaluate the clinical value of this scale.[16,17,18] These studies all indicated a higher sensitivity and specificity in screening for BD with the BI, and the patients with a positive family history, more frequent episodes, and a younger age at first onset, tend to get higher BI scores. However, these studies differed with respect to indexes of diagnostic value, such as sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV).[14,15,16,17,18]

BD is often not recognized, which can lead to potentially drastic consequences for the individuals and their families. Therefore, it is very important that a diagnostic scale is used in clinical practice. However, the study results of diagnostic scales for BD must been summarized and analyzed. Despite some limitations, using the HCL-32 as a first screening test in patients seeking help for depression can be recommended, but should never be used on its own for diagnosis based on meta-analysis.[19] It is also suggested that future research should examine whether screening properties can be improved by developing an algorithm incorporating the negative consequences reported for different areas in the HCL-32. A systematic review and meta-analysis of accuracy studies for the bipolar spectrum diagnostic scale (BSDS), the HCL-32, and the MDQ were also performed.[20] Although accuracy properties of the 3 screening tools did not consistently differ in mental health care services, the HCL-32 was more accurate than the MDQ for the detection of type II BD. It means that meta-analysis of diagnostic scales may find some new results. This study is only a meta-analysis of the data from past clinical studies in China, and analyzes the diagnostic value of the BI scale again.

METHODS

Literature Retrieval Methods

This study was performed according to the recommendations of the MOOSE (Meta-analysis Of Observational Studies in Epidemiology) guidelines.[21] Two reviewers independently searched the database, which included all Chinese databases: Chinese Biomedical Database (CBM), China National Knowledge Infrastructure (CNKI), WANFANG, and Chinese Social Sciences Citation Index (VIP).

Search key words: BI; BD (mood disorder, mania, bipolar depression, depression).

The Search Strategy: The search strategy was based on combinations. To retrieve all articles, we search papers by querying "BI and BD (or mood disorder or mania or bipolar depression or depression)," and then further screened the related papers. The last query was updated for the period

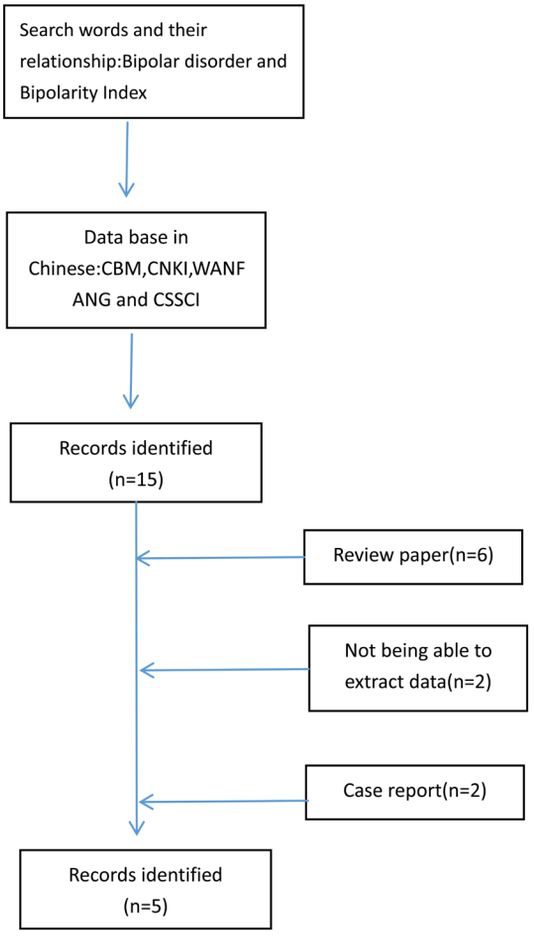

Figure 1. Flowchart of selection of studies for inclusion in meta-analysis.

Table 1. Characteristics of Studies Included in the Meta-analysis

Author	Cases	Participants	BI	Diagnostic Criteria and Other Assist Tool	Index	Quality Score
Ma[16]	727	BD, BD-II, MDE	BI(Chinese version)	DSM-IV-TR	Sensitivity Specificity Accuracy	8
Zhu[14]	95	BD, MDE	BI(Chinese version)	DSM-IV HCL-32	Sensitivity Specificity Accuracy	8
Guo[15]	176	BD, MDE	BI (Chinese version)	MINI, ADE	Sensitivity Specificity Accuracy PPV NPV	8
Li[17]	120	BD, MDE	BI (Chinese version)	DSM-5 ICD-10	Sensitivity Specificity Accuracy PPV NPV	8
He[18]	120	BD, RMDD	BI (Chinese version)	DSM-IV, MINI	Sensitivity Specificity Accuracy	8

BD, bipolar disorder; MDE, major depression episode; BI, bipolarity index; MINI, mini-international neuropsychitric interview; DSM-IV, American Diagnostic and Statistic Manual of Mental Disorders, 4th edition; RMDD, recurrent major depressive disorder; ADE, affective disorder evaluation.

between July 31, 2004 and July 31, 2020. References of retrieved articles were cross-searched to identify any studies missed by the electronic search (Figure 1).

Inclusion and Exclusion Criteria

The 2 researchers independently reviewed the initially retrieved publications. Discrepancies were resolved through discussion by all reviewers. Studies that met the following criteria were included: (1) study about BI in the diagnosis of BD, or (2) studies about BI to screen for diagnosis of BD in a clinical setting, or (3) study about BI used in the diagnosis of BD in patients with depression, or (4) study about BI used in the diagnosis of BD in patients affected by other disorders, and (5) study papers written in Chinese. However, articles that had incomplete or unidentified data were excluded, as well as abstracts, reviews, case reports, letters, and duplicate publications.

Two psychiatrists reviewed each included article independently, using the 11-item checklist that was recommended by the Agency for Healthcare Research and Quality (AHRQ).[22] Every article was coded according to the first letter of the first author. If 2 or more first authors had the same first letter, Arabic numerals were added to show the difference. The psychiatrists were trained in evidence-based medicine.The disputes were discussed and resolved according to the scale. An item would be scored “0” if it was answered “NO” or “UNCLEAR,” whereas “1” if the answer was “YES.” Article quality was assessed as follows: low quality , 0-3; moderate quality , 4-7; and high quality , 8-11. Differences in article quality were discussed to reach an agreeable final score. The following information was extracted: first author, date of publication, the sample size, study population, assessment tools, and the statistics on sensitivity, specificity, accuracy, PPV, and NPV (Table 1).

Statistical Analysis

All statistical analyses were performed using software of RevMan 5.2 and the *P* value for the overall effect < .05 with the 2-tailed test was considered statistically significant. The heterogeneity of all involved studies was assessed by the I^2 statistic. When it was lower than 50%, the studies with an acceptable heterogeneity were considered, and then the fixed-effects model with Mantel–Haenszel method was used; otherwise, a random effect model with the DerSimonian and Laird method was adopted.

Assessment of publication bias was investigated for each of the pooled study groups, mainly by Egger’s linear regression test. As a supplementary approach, Begg’s rank correlation also was applied to assess the potential publication bias.

RESULTS

A total of 5 studies, with 1237 patients, met the inclusion criteria and were included in the final meta-analysis. The 1237 patients were all classified as samples diagnosed with BD or without BD. The sample size of the studies ranged from 75 to 727. The revised Chinese version of the BI was the assessment tool used in the studies. [6] The main features of the 5 articles have been summarized in Table 2. The AHRQ scores suggested that all 5 studies scored 8 points, indicating high quality.

The random effect model was used to account for the data, with RevMan 5.2. The results showed that the

Table 2. The Primary Data of 5 Studies

	Sensitivity(n/N)	Specificity(n/N)	Accuracy(n/N)	PPV(n/N)	NPV(n/N)
Li(2015)	22/26	78/94	100/120	22/38	78/82
He(2014)	52/56	60/60	116/120	56/60	60/64
Guo(2014)	99/108	63/68	162/176	99/104	63/72
Ma(2013)	290/290	307/437	597/727	290/420	307/307
Zhu(2011)	17/20	47/75	64/95	17/45	47/50

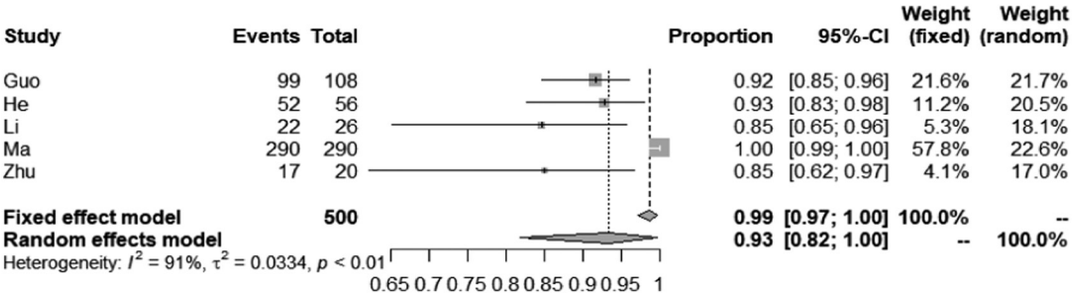

Figure 2. Pooled sensitivity of meta-analysis in 5 studies. Results showed that the sensitivity of BI in diagnosis for BD was 0.93 (95% CI: 0.93-1.00).

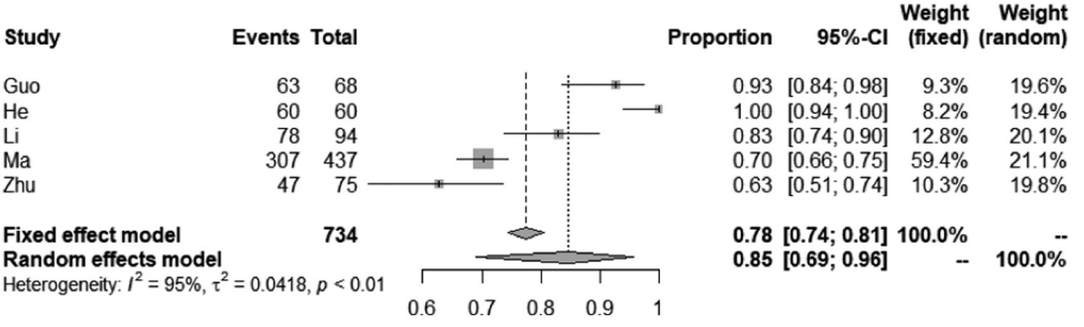

Figure 3. Pooled specificity of meta-analysis in 5 studies. The specificity was 0.85 (95% CI: 0.69-0.96).

sensitivity of BI in diagnosis for BD was 0.93 (95% CI: 0.93-1.00) (Figure 2). The specificity was 0.85 (95% CI: 0.69-0.96) (Figure 3). The accuracy was 0.86 (95% CI: 0.77-0.93) (Figure 4). The PPV was 0.74 (95% CI: 0.53-0.91) (Figure 5). The NPV was 0.95 (95% CI: 0.81-1.00) (Figure 6).

Significant heterogeneity was detected across studies regarding these incidence estimates, which all were higher than 90%.

Assessment of publication bias for each of the pooled study groups, mainly by Egger's linear regression test and Begg's rank correlation, was not investigated because of a sample size smaller than 9.

DISCUSSION

Rates of misdiagnosis between MDD and BD have been reported to be substantial, and the consequence of such

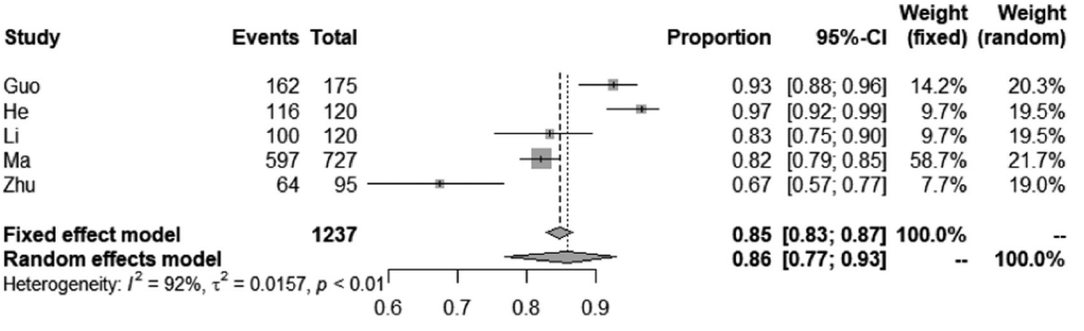

Figure 4. Pooled accuracy of meta-analysis in 5 studies. The accuracy was 0.86 (95% CI: 0.77-0.93).

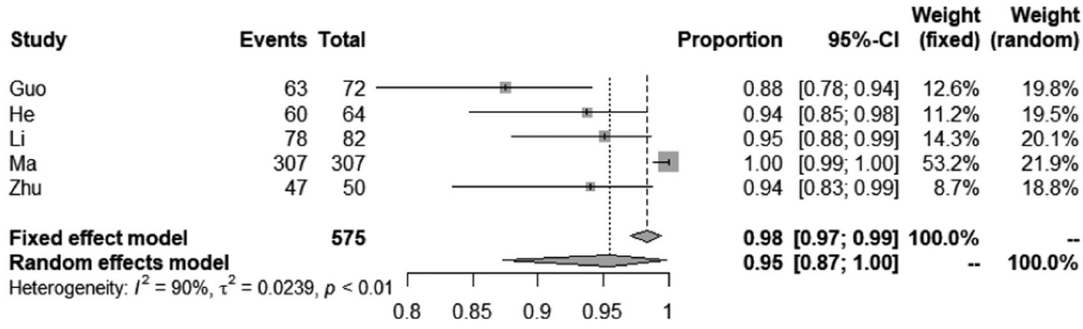

Figure 5. Pooled PPV of meta-analysis in 5 studies. The positive predictive value was 0.74 (95% CI: 0.53-0.91).

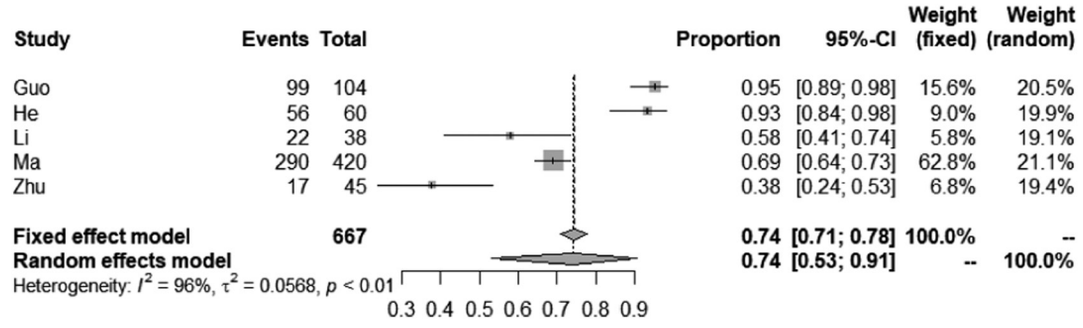

Figure 6. Pooled NPV of meta-analysis in 5 studies. The negative predictive value was 0.95 (95% CI: 0.81-1.00).

misdiagnosis is likely to be a delay in achieving effective control of symptoms, in some cases spanning many years. Particularly in the midst of a depressive episode, or early in the course of illness, it may be challenging to distinguish the 2 mood disorders purely on the basis of cross-sectional features. To date, no useful biological markers have been shown to reliably distinguish between BD and MDD.[23] Therefore, it may be feasible to diagnose BD by clinical phenomenology, which includes age of onset, symptoms, and response to treatment drugs such as mood stabilizers, atypical antipsychotics, or antidepressants.

Inspired by clinical experience and driven by an intent to assign a "bipolar profile" for the individual patient seeking treatment for a probable mood disorder, experienced investigators in the field of BD created the BI in 2004.[24] The BI is a 0-100 continuous scale that covers 5 illness dimensions with a maximum of 20 points per domain: I, signs and symptoms; II, age of onset; III, course of illness; IV, response to treatment; and V, family history. "Classic" BD, according to the authors of the BI, would be characterized by: I, at least one euphoric manic episode; II, early age of onset; III, recurrent, and fully remitting illness course; IV, positive response to a mood stabilizer; and V, having a first-degree family member with BD. The 5 dimensions of the BI highlight that the conceptualization of "bipolarity" that underlies the BI goes beyond the mere symptomatic assessment of lifetime affective symptoms, as the DSM-IV categorization requires. Instead, the authors of the BI added additional state and trait variables, based on their clinical experience, and on the earlier theories of Kraepelin.[7] By doing so, the BI represents a broader view of mood disorders that is now termed "bipolarity," that represents a more conservative view of the classic conceptualization of mood disorders, which Kraepelin considered to include risk for shifts to elevated mood states. Hence, it is well possible that the BI estimates a latent trait of bipolarity that may become apparent in bipolar conversion at a later stage in those with a lifetime unipolar disorder(UD). However, studies on the concurrent validity of the BI against a lifetime DSM-IV classification of BD are potentially flawed by observer bias,[25] as the diagnosis of BD and BI answers are often provided by the same clinician, with good to excellent metrics.[5,6,26] Therefore, comprehensive analysis of BI for diagnosis is needed, which also was the original intention of this study, although it was limited to China.

In China, many psychiatrists are interested in the application of BI in the diagnosis of BD, and have carried out a series of related studies. Most studies support the important diagnostic function of BI, although there are some different conclusions. Therefore, it is meaningful to evaluate the results comprehensively, which is the original intention of our study.[14-18] Diagnostic sensitivity, also known as true positive rate, refers to the probability of actual illness, and being diagnosed with illness. It reflects the ability to test patients. The greater the value, the better. Diagnostic specificity, also known as true negative rate, refers to the probability that a diagnostic test identifies individuals who are actually disease-free and are diagnosed as disease-free, which reflects the ability to

identify non-patients. The greater the value, the better. The focus of this study is on them.

These studies, with 1237 subjects, met the inclusion criteria, and were included for the final meta-analysis. The sample size of the studies ranged from 75 to 727. The assessment tool used in the studies is the revised Chinese version of the BI.[6] The result in this study found that the sensitivity of BI for diagnosis of BD was 0.93 (95% CI: 0.93-1.00), and the specificity was 0.85 (95% CI: 0.69-0.96). The PPV was 0.74 (95% CI: 0.53-0.91). The NPV was 0.95 (95% CI: 0.81-1.00), and accuracy was 0.86 (95% CI: 0.77-0.93). Based on the data of our study, BI was found to be the ideal tool of diagnosis for BD. In fact, the same results have also been obtained in other studies outside of China.[5,7,26] Significant heterogeneity was also found to be higher across studies regarding these incidence estimates, which means that we must interpret the results with caution, and pay attention to them.

Compared to other assistive diagnostic scales for diagnosis of BD, such as HCL-32, MDQ, and BSDS, the BI includes age of onset, family history, response to treatment, and course of illness. Mosolov`s study found that total BI score was significantly higher in the BD-II group than in the recurrent depressive disorder (RDD) group: 31.8 (SD = 12.4) versus 20.2 (SD = 7.8), respectively ($P < .0001$). It can significantly differentiate BD from RDD.[26] Aikens's study found that at a cut-off of ≥50, the BI had a high sensitivity (0.91) and specificity (0.90).[5] Ma's study found that the cut-off score between the MDD and BPD groups was 42.0, with a sensitivity of 0.957 and specificity of 0.881 ($Z = 63.064$, $P < .001$); the cut-off score between the MDD and BPD II groups was 34.0, with a sensitivity of 0.810 and specificity of 0.855 ($Z = 20.174$, $P < .001$); and the cut-off score between the BPD II and BPD I groups was 57.0, with a sensitivity of 0.680 and specificity of 0.772 ($Z = 9.636$, $P < .001$).[6] Wendela G. ter Meulen and their colleagues found each point increase in BI score significantly predicted incident BD (HR [95%CI] = 1.047[1.018-1.076], $P = .001$). At the optimal cut-off of 30, sensitivity was 67%, specificity 52%, PPV 3%, and NPV 98%.[7] These indicated that BI has a special role in diagnosis of BD and maybe more useful than HCL-32, and MDQ.[5,6,7,26]

The age of onset is significant in the BI. The younger the age of first onset, the higher the score on the scale, for example, the score is 20 when age of first onset was 15-19 years.[24] Younger age of first onset also indicates a higher probability of BD. Many current studies have discussed the diagnosis of BD in children throughout the years as it has evolved, focusing on very early-onset and early-onset BD. Proper care of children with BD requires a thorough understanding of the subtleties in symptoms at different developmental ages, as well as a shift in diagnostic thinking, which grew to include disruptive mood dysregulation disorder.[27] This also means age of onset may have an etiological value. In fact, in clinical practice, early age of onset was found to be associated with longer delay in treatment (Hedges' g = 0.39, $P = .001$), greater severity of depression (Hedges' g = 0.42, $P < .001$), higher levels of comorbid anxiety (OR = 2.34, $P < .001$), and substance use (OR = 1.80, $P < .001$).[28] So it is significant to note the age of onset, not only in terms of the BI, but also necessary in clinical psychiatry.

Recurrent distinct manic episodes separated by periods of full recovery were described in patients with a BI score of 20, which was a special course of illness, and comorbid substance abuse was given a score of 10, Recurrent unipolar MDD with 3 or more episodes of MDD was scored 5.[24] A study found that among patients with a current diagnosis of RDD, 40.8% had a diagnosis of bipolar disorder (bipolar I disorder: 4.9%; BD-II: 35.9%).[26] Certain risk factors such as the young age of onset and greater episode frequency are useful predictors of bipolar diatheses. Substance use disorder comorbidity is more prevalent in males whereas depression and suicidal behavior are more frequent in females with BD. Comorbid anxiety and personality disorders also encumber the illness course.[29] It means that focusing attention on the course of illness will not only help diagnosis of BD, but also understand the traits of BD.

Worsening dysphoria or mixed symptoms during antidepressant treatment subthreshold for mania were given a score of 10 in BI,[24] which actually indicated the use of antidepressant in bipolar depression and mixed episodes. It is a very controversial issue in psychopharmacology. The study found the rate of manic switch in AD-m (antidepressant monotherapy) was significantly higher than the AD-c (antidepressant combination) group.[30] It suggests that the risk of manic switch is especially prominent in the first months of AD use. Antidepressants used in combination with mood stabilizers (MS) may not be adequate in preventing switches in the short-term. However, in the longer term, addition of MS to ADs may decrease the risk of switches. The correct selection of adjunctive second-generation antidepressant therapy with a mood stabilizer or an atypical antipsychotic is important in treatment for bipolar depression.[31] The key point is switching to mania in the treatment of unipolar depression, which had been solved in DSM-5 that can be diagnosed as BD.[32] Antidepressant-induced mania or hypomania in first edition of soft bipolar criteria was a very important item,[11] but was revised out of the second criteria of soft bipolar with the publication of DSM-5.[13]

The score of 20 in the items of the BI scale was assessed as "at least 1 first-degree relative with documented bipolar illness"; a score of 10 was assessed as "first-degree relative with documented recurrent unipolar MDD or schizoaffective disorder or any relative with documented bipolar illness or any relative with documented recurrent unipolar MDD and behavioral evidence suggesting bipolar illness."[24] This suggests that BD has a heavier genetic load. More and more evidence supports the fact that a significant relationship

exists between the degree of kinship and the heritability of BD. In fact, a positive family history has a higher diagnostic specificity for BD,[13,32] A dimensional definition based on 3 or more hypomanic symptoms during depression was the most supported by using bipolar family history as validator.[32] At same time, patients reported family history of a mood disorder had an earlier age at onset of depression/mania, more phases, rapid cycling, and more frequent suicide attempts. Across different assessments, patients with family history showed consistently elevated depressive symptoms, such as lower concentration and energy, higher suicidal thoughts, as well as increased racing thoughts and distractibility within the manic spectrum of symptoms.[33]

This is the first study to analyze the application of BI in the diagnosis of BD, suggesting that BI is an important auxiliary diagnostic scale in the diagnosis of BD. Although the BSDS has completed the screening of bipolar spectrum disorders,[20] it may contain more information than the MDQ and HCL-32. Fortunately, this study suggests that BI has better sensitivity and specificity.

This study had several limitations. Firstly, the sample size of this meta-analysis was relatively small. Only 5 studies and 1237 subjects were involved. It is difficult to reflect the BI value in clinical psychiatry in China. Secondly, the method of data collection may influence the result of investigation; for example, different cut-off criteria can result in different detection rates of sensitivity, specificity, PPV, NPV, and accuracy. So it was very important to establish a standard cut-off or sub-analysis according the different cut-off criteria. Thirdly, not all the studies had the same diagnostic criteria for BD, which maybe a significant factor affecting sensitivity, specificity, PPV, NPV, and accuracy. Fourthly, it is notable that studies included in this meta-analysis had high heterogeneity and also had certain bias in data, which cannot been reflected by Egger's linear regression test and Begg's rank correlation because of the small sample size. These factors are partly responsible for the prospective heterogeneity of sensitivity, specificity, PPV, NPV, and accuracy. It also affects the representation of the real significance of BI in clinical psychiatry, and its difference from other assistive diagnostic scales.

CONCLUSION

A total of 1237 subjects were included in 5 studies. The random effect model was used to account for the data with RevMan 5.2. The results showed that the diagnostic sensitivity of BI in was 0.93 (95% CI: 0.93-1.00), the specificity was 85% (95% CI: 0.690.96). The PPV was 74% (95% CI: 0.53-0.91). The NPV was 95% (95% CI: 0.81-1.00), and accuracy was 86% (95% CI: 0.77-0.93). Significant heterogeneity was detected across studies regarding these incidence estimates. The ideal diagnostic value of BI was found. Despite the significant heterogeneity detected in studies, we must interpret the results with caution and attention, which includes comparison to other diagnostic scales in perfecting the use of BI in clinical psychiatry. It should be noted that personality and temperament may also be important factors that are not involved in BI.

Ethics Committee Approval: N/A.

Informed Consent: N/A.

Peer Review: Externally peer-reviewed.

Author contributions: Concept – S.F., J.W.; Design - J.W.; Supervision - J.W.; Resource - S.F.; Materials - T.H., S.F.; Data Collection and/or Processing - Z.J.; Analysis and/or Interpretation - S.F., J.W., Literature Search - S.F., T.J.; Writing - S.F.; Critical Reviews - J.W.

Acknowledgments: We thank Prof. Liu Tiebang Baopang (Shenzhun Mental Health Center) for his suggestion on the study concept and Mr Ren Zhibin for his help in literature retrieval. We thank Ms Ren Xin for her help in part of the statistical analysis. We thank Prof Fang Marong for the final revision of the article.

Conflict of interest: The authors have no conflicts of interest to declare.

Financial disclosure: The authors declared that this study has received no financial support.

REFERENCES

1. Stiles BM, Fish AF, Vandermause R, Malik A. Identifying the complexity of diagnosing bipolar disorder: A focused ethnography. *Issues Ment Health Nurs*. 2019 ;40(9):812-818. [CrossRef]
2. Fritz K, Russell AMT, Allwang C et al. Is a delay in the diagnosis of bipolar disorder inevitable?*Bipolar Disord*. 2017;19(5):396-400. [CrossRef]
3. Lee K, Oh H, Lee EH et al. Investigation of the clinical utility of the hypomania checklist 32 (HCL-32) for the screening of bipolar disorders in the non-clinical adult population. *BMC Psychiatry*. 2016;16:124. [CrossRef]
4. Feng Y, Wang YY, Huang W et al. Comparison of the 32-item Hypomania Checklist, the 33-item Hypomania Checklist, and the Mood Disorders Questionnaire for bipolar disorder. *Psychiatry Clin Neurosci*. 2017 ;71(6):403-408. [CrossRef]
5. Aiken CB, Weisler RH, Sachs GS. The bipolarity index: A clinician-rated measure of diagnostic confidence. *J Affect Disord*. 2015;177:59-64. [CrossRef]
6. Ma Y, Gao H, Yu X et al. Bipolar diagnosis in China: evaluating diagnostic confidence using the bipolarity index. *J Affect Disord*. 2016;202:247-253. [CrossRef]
7. Ter M, Wendela G, Draisma S, Beekman ATF, Penninx BWJH, Kupka RW. The predictive performance of the bipolarity index in a Dutch epidemiological sample manuscript. *J Affect Disord*. 2020;262:373-380. [CrossRef]
8. Thomas RK, Baker G, Lind J, Dursun S. Rapid effectiveness of intravenous ketamine for ultraresistant depression in a clinical setting and evidence for baseline anhedonia

and bipolarity as clinical predictors of effectiveness. *J Psychopharmacol*. 2018 ;32(10):1110-1117. [CrossRef]
9. Wang YM, Song M, Shi SX et al. Relationship between affective temperament and bipolarity in depressive episode. J Neurosci and Mental Health. 2015.04.018;15(4):386-388. [CrossRef]
10. Li ZY, Yu X, Ma YT, Dang WM. Investigation of applicability of bipolarity index by Delphi method. *Chin J Nerv Ment Dis*. 2014;40(2):102-105. [CrossRef]
11. Jin WD. Soft bipolar depression progress in China. *Open J Depress*. 2015;04(4) :53-57. [CrossRef]
12. Kuppili PP, Yadav P, Pattanayak RD. Concept and Identification of "Soft Bipolarity" in Patients Presenting With Depression: Need for Careful Screening by Physicians. *J Assoc Physicians India*. 2017 ;65(7):80-84. PubMed: [CrossRef].
13. Li BL, Gao ZH, Jin WD. Research progress of diagnostic criteria of soft bipolar disorder. *J Clin Psychosom Dis*. 2016;22(3):113-115. [CrossRef]
14. Zhu Y, Ma YT, Wei J, Yu X. Recognition potency of three diagnostic definition of bipolar disorder in patients with current depressive episode. *Chin Ment Health J*. 2011;25(8):588-593. [CrossRef]
15. Guo XB, Xue HX, Feng L, Geng Y Study of reliability and validity of Chinese version of the bipolarity index in clinic. *China Med Herald*. 2014;24(11):119-123.
16. Ma YT, Yu X, Wei J, Zheng Y. Recognition validity of bipolarity specifier for bipolar disorders among patients with major depressive episode:BRIDGE-China.*Chin J Psychiatry*. 2013;46(5):271-276. [CrossRef]
17. Li F. Comparative study on efficacy of identification between bipolarity index and DSM-V criteria in screening patients with bipolar disorder. *Chin Mod Med*. 2015;22(19):66-69.
18. He HZ, Sun j, Zhu RX, Cheng WR. Recognition performance of bipolar index on bipolar disorder. *J Clin Psychiatry*. 2014;24(1):8-10.
19. Meyer TD, Schrader J, Ridley M, Lex C. The hypomania checklist (HCL) - systematic review of its properties to screen for bipolar disorders. *Compr Psychiatry*. 2014 ;55(5):1310-1321. [CrossRef]
20. Carvalho AF, Takwoingi Y, Sales PMG et al. Screening for bipolar spectrum disorders: A comprehensive meta-analysis of accuracy studies. *J Affect Disord*. 2015;172(172):337-346. [CrossRef]
21. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000;283(15):2008-2012. [CrossRef]
22. *Celiac disease*. Rockville: Agency for Healthcare Research and Quality; 2004
23. Vöhringer PA, Perlis RH. Discriminating Between bipolar disorder and major depressive disorder. *Psychiatr Clin North Am*. 2016 ;39(1):1-10. [CrossRef]
24. Sachs GS. Strategies for improving treatment of bipolar disorder: Integration of measurement and management. *Acta Psychiatr Scand Suppl*. 2004;110(422):7-17. [CrossRef]
25. Holman L., Head ML, Lanfear R., Jennions MD. Evidence of experimental bias in the life sciences: why we need blind data recording. *PLOS Biol*. 2015;13(7):e1002190. [CrossRef]
26. Mosolov S, Ushkalova A, Kostukova E et al. Bipolar II disorder in patients with a current diagnosis of recurrent depression. *Bipolar Disord*. 2014;16(4):389-399. [CrossRef]
27. Roselle A.Pediatric bipolar disorder: onset, risk factors, and protective factors. *J Psychosoc Nurs Ment Health Serv*. 2019;57(9):32-37. [CrossRef]
28. Joslyn C. , Hawes D. J , Hunt C. , Mitchell P B .Is age of onset associated with severity, prognosis, and clinical features in bipolar disorder? A meta-analytic review.*Bipolar Disord*. 2016 ;18(5):389-403.[CrossRef]
29. Subramanian K, Sarkar S, Kattimani S. Bipolar disorder in Asia: Illness course and contributing factors. *Asian J Psychiatr*. 2017;29:16-29. [CrossRef]
30. Berkol TD, Balcioglu YH, Kirlioglu SS et al. Clinical characteristics of antidepressant use and related manic switch in bipolar disorder. *Neurosciences (Riyadh)*. 2019 ;24(1):45-52. [CrossRef]
31. McGirr A, Vöhringer PA, Ghaemi SN, Lam RW, Yatham LN. Safety and efficacy of adjunctive second-generation antidepressant therapy with a mood stabiliser or an atypical antipsychotic in acute bipolar depression: A systematic review and meta-analysis of randomised placebo-controlled trials. *Lancet Psychiatry*. 2016 ;3(12):1138-1146. [CrossRef]
32. Terao T, Tanaka T. Antidepressant-induced mania or hypomania in DSM-5. *Psychopharmacol (Berl)*. 2014;231(1):315. [CrossRef]
33. Benazzi F. Family history validation of a definition of mixed depression. *Compr Psychiatry*. 2005;46(3):159-166. [CrossRef]