

Published in final edited form as:

Clin Psychol (New York). 2009 June 1; 16(2): 188–201. doi:10.1111/j.1468-2850.2009.01158.x.

Assessment Tools for Adult Bipolar Disorder

Christopher J. Miller, Sheri L. Johnson, and Lori Eisner

University of Miami

Abstract

This article reviews the current state of the literature on the assessment of bipolar disorder in adults. Research on reliable and valid measures for bipolar disorder has unfortunately lagged behind assessment research for other disorders, such as major depression. We review diagnostic tools, self-report measures to facilitate screening for bipolar diagnoses, and symptom severity measures. We briefly review other assessment domains, including measures designed to facilitate self-monitoring of symptoms. We highlight particular gaps in the field, including an absence of research on the reliable diagnosis of bipolar II and milder forms of disorder, a lack of empirical data on the best ways to integrate data from multiple domains, and a shortage of measures targeting a broader set of illness-related constructs relevant to bipolar disorder.

Keywords

adult; assessment; bipolar; diagnosis; screening; severity

The goal of this review is to summarize measures that are useful for the assessment of bipolar disorder among adults. We will focus, in particular, on measures pertinent to screening, diagnosis, and symptom monitoring. With the apparent success of lithium in treating bipolar disorder, research on the disorder languished until the 1990s. Interest in bipolar disorder assessment has been renewed in recent decades. Nonetheless, research on the accurate assessment of bipolar disorder is relatively sparse when compared with other disorders such as major depression. We begin by describing the forms of bipolar disorder, then turn to available measures for its diagnosis, including both interview and self-report measures. Later sections discuss interviews and scales used for assessing symptom severity, including self-monitoring.

NATURE OF BIPOLAR DISORDER

Several types of bipolar disorder are recognized by the *Diagnostic and Statistical Manual of Mental Disorders* of the American Psychiatric Association (APA, 2000), differentiated by the severity and duration of manic symptoms. A diagnosis of bipolar I disorder is made based on a single lifetime episode of mania, which is in turn defined by euphoric or irritable mood, along with at least three additional symptoms (or four if mood is only irritable) that result in marked social or vocational impairment. The duration criterion for mania specifies that symptoms must last one week or require hospitalization. Bipolar II disorder, in contrast, is defined by a history of at least one hypomanic episode and at least one major depressive episode. Criteria for hypomania are similar to those of mania, but in milder form: instead of impairment, a hypomanic episode is marked by a distinct change in functioning. Cyclothymic disorder is an even milder subtype of bipolar disorder, and is diagnosed based on a period of at least two

years of recurrent mood swings. By definition, these mood swings must be in both the “up” and the “down” directions, but do not meet full criteria for mania, hypomania, or depression. In addition, the symptomatic two-year period cannot include any two-month span that is free of mood swings.

Symptoms that are secondary to drugs such as cocaine, or medical conditions such as thyroid problems, will generally yield a diagnosis of substance-induced mood disorder or bipolar disorder not otherwise specified. Those with a vulnerability to bipolar disorder may become manic when prescribed antidepressants without an accompanying mood stabilizer (Ghaemi, Lenox, & Baldessarini, 2001), yielding a diagnosis of substance-induced mood disorder with manic features.

Large epidemiological studies indicate a prevalence of 1% for bipolar I disorder and an additional 3% for bipolar II disorder (Kessler, Berglund, Demler, Jin, & Walters, 2005). As many as three quarters of those with bipolar I disorder have also experienced an episode of major depression (Karkowski & Kendler, 1997; Kessler, Rubinow, Holmes, Abelson, & Zhao, 1997). Comorbidity rates with anxiety disorders and substance abuse disorders have been reported as high as 93% and 61%, respectively (Kessler et al., 1997; Regier et al., 1990), underscoring the need for effective assessments and treatments of bipolar disorder to take comorbid conditions into account.

Twin studies suggest that heritability accounts for more than 90% of the variability in the development of bipolar disorder (Kieseppä, Partonen, Haukka, Kaprio, & Lonnqvist, 2004), leading many researchers to focus on medications such as lithium for treatment (Prien & Potter, 1990). The course of the disorder, however, may be strongly affected by psychosocial variables. Manic episodes may be triggered by sleep disturbance (Leibenluft et al., 1996) or excessive pursuit of goals (Johnson, 2005). Depressive episodes within bipolar disorder share common triggers with unipolar depression, such as negative life events, maladaptive cognitive styles, and lack of social support (Johnson & Kizer, 2002). Thus psychotherapy may serve as an effective addition to medication in the treatment of bipolar disorder (Johnson & Leahy, 2004; Rizvi & Zaretsky, 2007).

ASSESSMENT TOOLS FOR DIAGNOSIS

The diagnosis of bipolar disorder is based on a review of symptoms and potential medical explanations for those symptoms, as there is no biological marker for the disorder. In clinical practice, symptoms are frequently reviewed in an unstructured manner. It should be noted, though, that when practitioners do not use structured diagnostic tools, as many as half of comorbid conditions go undetected (Zimmerman & Mattia, 1999). Furthermore, many practitioners report that they do not routinely screen for bipolar disorder even among people with a history of major depression, many of whom would meet the diagnostic criteria for bipolar disorder (Brickman, LoPiccolo, & Johnson, 2002). Due to informal or poor screening, the average time between onset of symptoms and formal diagnosis is more than seven years (Lish, Dime-Meenan, Whybrow, Price, & Hirschfeld, 1994; Mantere, Suominen, Leppamäki, Arvilommi, & Isometsä, 2004). Improper diagnosis has serious repercussions because antidepressant treatment without mood-stabilizing medication can trigger iatrogenic mania (Ghaemi et al., 2001).

Several semistructured interviews have been developed to assess bipolar disorder in adults. The two most commonly used measures are the Structured Clinical Interview for *DSM-IV* (SCID) and the Schedule for Affective Disorders and Schizophrenia (SADS). We will not focus here on the Composite Interview Diagnostic Interview (CIDI; Robbins et al., 1988), which has been developed and used mostly in epidemiological surveys (e.g., Kessler & Zhao, 1999). Briefly, there is some evidence that the CIDI may systematically underdiagnose bipolar

disorder (e.g., Kessler, Rubinow, Holmes, Abelson, & Zhao, 1997), but more recent work has since validated it against the SCID (Kessler et al., 2006). The SCID and the SADS both provide interview probes, symptom thresholds, and information about exclusion criteria (i.e., medical or pharmacological conditions that may induce mania). They differ, however, in the criteria they were designed to assess. The SCID is designed to help assess diagnoses according to the *DSM-IV*, whereas the SADS is designed to assess diagnoses according to the Research Diagnostic Criteria (RDC). RDC criteria are stricter in that psychotic symptoms are more likely to yield a diagnosis of schizoaffective disorder than would be applied in the *DSM-IV* criteria; within the *DSM-IV* criteria, psychotic symptoms must be present for at least two weeks outside of episode to be considered evidence of schizoaffective disorder. Further details about these measures are provided next. We begin by describing the measures and their psychometric characteristics for assessing bipolar I disorder. We then turn toward some specific issues that complicate the assessment of milder forms of bipolar disorder. Table 1 summarizes some of the well-supported measures for the diagnosis of bipolar disorder.

The SCID (Spitzer, Williams, Gibbon, & First, 1992) is recommended as a routine part of clinical intake procedures. The SCID is a semistructured interview that is divided into modules to cover different diagnoses. The modular design allows for the interview to be easily tailored to capture relevant diagnoses for a given research or clinical situation. Each SCID module contains probes to cover each of the core symptoms, and interviewers can use clinical judgment in gathering supplemental information if probes do not provide sufficient information for reliable symptom assessment. A clinician's version is available through American Psychiatric Publishing (First, Spitzer, Gibbon, & Williams, 1997). The SCID, and more specifically its bipolar disorder module, demonstrated good interrater reliability both in a large international multisite trial (Williams et al., 1992) and in at least 10 other major trials (Rogers, Jackson, & Cashel, 2001). In patient samples, reliability for current and lifetime diagnoses of bipolar disorder has been adequate to excellent, ranging from .64 to .92; establishing reliability for the SCID in community samples is more difficult due to low base rates of the disorder (Williams et al., 1992). Compared to other structured interviews including the Diagnostic Interview Schedule (DIS) and the Composite International Diagnostic Interview (CIDI), and to clinicians not using a structured interview, diagnoses of bipolar disorder based on the SCID appear substantially more reliable. Results of one study indicated that the percentage of agreements with the gold standard were higher for the SCID as compared to standard clinician interviews (Basco et al., 2000). In a sample of twins, diagnoses of bipolar disorder made using the SCID showed similar concordance rates between monozygotic and dizygotic twins compared to traditional twin studies using standard diagnostic interviews (Kieseppä, et al., 2004).

The SADS (Endicott & Spitzer, 1978) was designed to assess a broad range of Axis I diagnoses. For each diagnosis, the probes focus on the symptoms for the most recent episode and then capture a broad overview of past episodes. The reliability and validity of the SADS has been established across 21 studies (see Rogers, Jackson, & Cashel, 2001, for a review). The SADS has demonstrated good to excellent reliability for both symptoms and diagnoses (Andreasen et al., 1981). Specifically, mania diagnoses have achieved good interrater reliability and achieved good test-retest reliability over 5 to 10 years among adults (Coryell et al., 1995; Rice et al., 1986). SADS diagnoses of bipolar disorder correlate robustly with other measures of mania (Secunda et al., 1985), and the SADS appears to validly capture diagnoses across different cultural and ethnic groups within the United States (Vernon & Roberts, 1982).

Diagnostic Assessment of Bipolar II Disorder in Adults

Hypomania is unique among *DSM* syndromes, in that by definition it does not cause any functional impairment. Perhaps because of this quality, the presence of at least one major depressive episode is also required to achieve a diagnosis of bipolar II disorder. This presents

a unique diagnostic challenge: the hypomanic episodes that separate bipolar II disorder from unipolar depression are by definition of only limited severity, making this a hard diagnosis to reliably detect. Complicating this picture is the fact that there are important disagreements in the field regarding the best criteria for hypomanic episodes. For instance, current *DSM* criteria require three or four symptoms, in addition to elevated or irritable mood, lasting at least four days. In contrast, RDC criteria only require three symptoms lasting two days. Given this uncertainty and relative lack of severity of hypomania, it is not surprising that the accurate assessment of bipolar II disorder is more difficult to achieve than bipolar I disorder.

Given that hypomania is almost always accompanied by less distress than depressive episodes, one might be tempted to focus on detecting depression. There is evidence, however, that the diagnosis of hypomania (and hence, bipolar II disorder) is important above and beyond the detection of depression. Diagnoses of bipolar II disorder are accompanied by increased mood lability (Akiskal et al., 1995) and a family history of bipolar II disorder (Rice et al., 1986). In addition, at least three studies have demonstrated that people with bipolar II disorder are at a higher risk for suicide than are those with bipolar I disorder or unipolar depression (Dunner, 1996). It is possible that the low mood of depression, combined with the impulsivity of hypomania, may be especially likely to lead to suicide attempts. In addition to suicide risk, the misdiagnosis of bipolar II disorder can have harmful pharmacological implications. The prescription of antidepressants, which is likely if bipolar II disorder is misdiagnosed as unipolar depression, may cause or exacerbate manic symptoms (Ghaemi et al., 2001). Thus, identification of bipolar II disorder may be pivotal in administering effective treatments.

The above-described difficulties in assessing hypomanic symptoms have manifested in low reliability for the SADS in detecting bipolar II disorder (Andreasen et al., 1981), even when interviewers rate the same tapes (Keller et al., 1981). Some research groups have achieved better estimates, however (Simpson et al., 2002; Spitzer & Endicott, 1978). Beyond the inconsistent estimates of interrater reliability, test-retest reliability over six months to two years likewise has been low for bipolar II disorder and cyclothymic disorder alike (Andreasen et al.; Rice et al., 1986). In one study, only 40% of participants with bipolar II disorder according to the SADS at baseline experienced any manic or hypomanic episodes over the ensuing 10 years (Coryell et al., 1995). This lack of ability to accurately detect bipolar II disorder is not limited to the SADS. In one study, a SCID interview missed one third of bipolar II cases identified by expert clinical interview (Dunner & Tay, 1993; Simpson et al., 2002). In sum, the best available diagnostic interviews are limited in their psychometric characteristics for the diagnosis of bipolar II disorder.

These difficulties have led some researchers to suggest that interviews aimed at detecting bipolar II disorder should start with questions about behavioral activation and increases in goal-directed behaviors rather than mood (Akiskal & Benazzi, 2005). Although promising, such approaches have not yet been fully validated.

In sum, a set of issues mars diagnosis of bipolar II disorder. Persons who meet criteria for bipolar II disorder may be at high risk for suicidality, and they may experience a worsening of manic symptoms if prescribed antidepressants. On the other hand, available tools do not detect bipolar II disorder reliably. Thus a major goal for ongoing research is to develop ways to reliably capture diagnoses of bipolar II disorder.

Self Report Measures

The most reliable and valid way to obtain a diagnosis of bipolar disorder is through a structured interview with a trained clinician (Akiskal, 2002). Nonetheless, given the time commitment involved in conducting structured interviews, several self-report measures have been developed to help clinicians identify persons most likely to meet criteria for bipolar disorders.

It should be emphasized that these measures do not provide diagnostic accuracy, but, rather, might help identify people who should warrant more careful diagnostic interviews.

The General Behavior Inventory (GBI) was designed to cover the core symptoms of bipolar disorder, including both depressive and manic symptoms (Depue et al., 1981). Different versions range from 52 to 73 items (e.g., Depue et al., 1981; Depue & Klein, 1988; Mallon, Klein, Bornstein, & Slater, 1986). Items on each version assess symptom intensity, duration, and frequency on a scale ranging from 1 (“never or hardly ever”) to 4 (“very often or almost constantly”). Although the GBI has the most robust psychometric properties of the available self-report screeners, the multiple versions make generalizations regarding psychometric properties difficult.

The full 73-item version of the GBI has demonstrated excellent internal consistency and adequate test–retest reliability. It has demonstrated sensitivity to bipolar disorder of approximately 75% and specificity greater than 97% (Depue & Klein, 1988; Depue et al., 1989; Klein, Dickstein, Taylor, & Harding, 1989; Mallon et al., 1986) in clinical and nonclinical samples. Cutoff scores, however, have not been consistent across studies, further limiting the generalizability of the scale. At present, the GBI appears to be a useful screening tool for bipolar disorder, but future research to establish norms and cutoffs would increase its utility.

Another screening tool is the Mood Disorder Questionnaire (MDQ; Hirschfeld et al., 2000). The first 13 items of the MDQ ask about the *DSM-IV* manic symptoms using a yes–no format. To achieve a positive screen, seven items must be endorsed. Additional items assess if the identified symptoms co-occurred and caused at least moderate impairment. The MDQ has attained adequate internal consistency (Hirschfeld et al., 2000; Isometsä et al., 2003), fair one-month test–retest reliability, and fair sensitivity (.73 to .90) in distinguishing between bipolar and unipolar disorder in clinical samples (Weber Rouget et al., 2005). In addition, at least one recent study has demonstrated that high MDQ scores are associated with greater impairment and suicidal ideation in a primary care setting (Das et al., 2005). Nonetheless, specificity has been low in some studies (.47 to .90; Hirschfeld et al., 2000, 2003; Isometsä et al., 2003; Miller et al., 2004; Weber Rouget et al., 2005) and the sensitivity in a community sample was only .28 (Hirschfeld et al., 2003).

A review of the content of MDQ items may help clarify why the scale has achieved better performance in inpatient settings than in community settings. Several of the items appear to capture common experiences in community samples. For example, in one study, as many as 90% of college students endorsed items such as “Have you ever had a time when you were not your usual self and you felt much more self-confident than usual?” (Miller, Johnson, & Carver, 2008). These items may be less commonly endorsed by persons with schizophrenia and other severe psychopathology, explaining why the scale may appear more beneficial in an inpatient setting than in a community sampling. Hence, the MDQ may be a potentially useful tool in clinical settings to screen for bipolar disorder among those with severe psychopathology, but may be less helpful in community settings.

Other scales appear helpful in nonclinical samples, but do not have enough data regarding their usefulness as screening tools in clinical settings. The Hypomanic Personality Scale (HPS; Eckblad & Chapman, 1986) predicted the development of manic episodes at 13-year follow-up in undergraduates (Kwapil et al., 2000). To date, the HPS has only been studied in one clinical sample, achieving a positive predictive value of .82 and a negative predictive value of .67, and achieving a point-biserial correlation of .56 with bipolar I diagnosis (Kwapil, 2008). The Bipolar Spectrum Diagnostic Scale (Ghaemi et al., 2005) and the Mood Spectrum Self-Reports (Dell’Osso et al., 2002) have only been examined in a single study each, and two

Hypomania Checklists (Angst et al., 2005; Hantouche et al., 2006) have only been examined in Europe and China (e.g., Meyer et al., 2007; Vieta et al., 2007). The Temperament Evaluation of Memphis, Pisa, Paris, and San Diego—Autoquestionnaire version (TEMPS-A; Akiskal & Akiskal, 2005) is a measure of temperament rather than manic or hypomanic episodes per se. Although the four-factor structure that includes dysthymic, cyclothymic, hyperthymic, and irritable temperaments has been examined in several countries and languages and psychometrically validated in clinical populations, research has not directly established the usefulness of this measure as a screen for bipolar spectrum disorders (e.g., Akiskal et al., 2005; Karam et al., 2007; Kesebir et al., 2005; Matsumoto et al., 2005; Mendlowicz, Jean-Louis, Kelsoe, & Akiskal, 2005; Sandor et al., 2006; Vazquez et al., 2007). At least one study, however, has demonstrated that the cyclothymic subscale of the TEMPS-A can prospectively predict bipolar spectrum diagnoses among clinically depressed children and adolescents over a two-year period (Kochman et al., 2005). Although initial studies indicate that these scales demonstrate good psychometric properties, more research is needed to determine their usefulness as screening measures.

Summary of Assessment Tools for Diagnosis

Overall, the SCID and the SADS are the most common means of diagnosing bipolar disorder in adults. With excellent psychometric characteristics for the assessment of bipolar I disorder, they fare less well in assessing bipolar II disorder. This may be due to issues related to the definition of hypomania.

As a diagnostic screening tool, the scale with the best support is the GBI, as it has consistently demonstrated sensitivity of approximately .75 and specificity above .97. Readers should be cautious, however, because multiple versions of the scale exist, and cutoffs for a positive screen have not been firmly established. The MDQ has been helpful in clinical populations, but suffers from poor discriminatory power in community settings. Other promising scales require more psychometric development. When using self-report scales as screening tools, several broader issues must be kept in mind. First, the usefulness of a screening tool will vary depending on the prevalence of a disorder in the population of interest (Phelps & Ghaemi, 2006). Second, few studies provide direct comparisons of psychometric characteristics of the different measures. Third, there are several ways to report on a screener's usefulness, including sensitivity and specificity, positive and negative predictive values, area under the curve, and point-biserial correlations with diagnosis (Kraemer, 1992). Not all studies on the detection of bipolar disorder report all of these results, limiting the ability to compare studies or measures. Furthermore, sensitivity and specificity are commonly reported, but these indices may be dependent on sample characteristics. Fourth, authors have often modified the diagnostic interviews used as a reference standard to capture milder forms of bipolar spectrum disorder, yet limited information about these modifications is available. Each of these issues makes comparisons between measures complex.

ASSESSMENT OF SYMPTOM SEVERITY

The most common approach to measuring the severity of manic symptoms has been clinician-rated interviews. The Young Mania Rating Scale (YMRS) and Bech-Rafaelsen Mania Rating Scale (MAS) are two of the most widely used clinician-rated scales for assessing symptom severity. These scales have been commonly used to track changes in symptoms over time as treatment progresses. We briefly review these two scales, as well as the Schedule for Affective Disorders and Schizophrenia—Change version (SADS-C) mania subscale. There has been growing recognition, though, of the need to track both clinician and patient perspectives on the course of treatment, and so we discuss available symptom severity measures that rely on self-report. Some research has focused on measures useful for case conceptualization and treatment planning, but this literature is not covered in detail here: interested readers are referred to other

reviews (e.g., Johnson, Miller, & Eisner, 2008). Table 2 summarizes some of the well-supported measures for assessing symptom severity in bipolar disorder.

The YMRS (Young, Biggs, Ziegler, & Meyer, 1978) is a 15- to 30-min interview designed to be conducted by a trained clinician. It was originally developed and tested within an inpatient population based on semi-structured interview and observation during an eight-hour period. Today, the YMRS combines the patient's report of manic symptoms over the previous two days as well as the clinician's observations during the interview. It consists of 11 items covering the "core symptoms of the manic phase": mood, motor activity, interest in sex, sleep, irritability, speech, flight of ideas, grandiosity, aggressive behavior, appearance, and an item regarding patient insight (Carlson & Goodwin, 1973; Winokur, Clayton, & Reich, 1969). It should be noted that item 8, Bizarre Content, combines the assessment of the manic symptom of grandiosity with other psychotic symptoms, including hyperreligiosity, paranoia, ideas of reference, delusions, and hallucinations. The YMRS does not account for other *DSM* criteria of mania, including distractibility, increases in goal-directed activity, or excessive involvement in pleasurable activities with a high potential for painful consequences. A factor analysis of the YMRS revealed a thought disturbance factor, an overactive/aggressive behavior factor, and a factor tapping elevated mood and psychomotor symptoms (Double, 1990).

Seven items are rated on a severity scale ranging from 0 to 4, and four items are rated on a scale of 0 to 8. Four core symptoms (irritability, speech, bizarre content, and disruptive–aggressive behavior) are double-weighted to account for poor cooperation from severely ill patients. Although the weighting may make rating more complex, it has not been shown to affect the reliability, validity, or sensitivity of the scale. The YMRS has demonstrated excellent psychometric properties, including a high inter-rater reliability for total scores (intraclass correlation = .93) and for individual item scores (intraclass correlation = .66 to .92), as well as high correlations with other mania rating scales (Young et al., 1978). Scores also statistically differentiate patients before and after two weeks of treatment. The YMRS has primarily been used to assess manic symptoms in treatment trials and was the primary measure of mania in the Systematic Treatment Enhancement Program for Bipolar Disorder study, the largest study to date on the effectiveness of treatments for bipolar disorder (Sachs et al., 2003).

The MAS (Bech et al., 1979) is a clinician-rated instrument that is similar in format to the YMRS. The 11 items of the MAS are rated on a five-point scale (ranging from 0 "not present" to 4 "severe") and cover classic manic symptoms such as elevated mood, irritability, sleep, increased activity, talkativeness, flight of ideas, self-esteem, noise level, and sexual interest. Like the YMRS, it has achieved excellent internal consistency and interrater reliability, as well as strong correlations with more exhaustive measures of manic symptoms (Bech, 1988; Bech, Bolwig, Kramp, & Rafaelsen, 1979; Licht & Jensen, 1997). It has been widely used in treatment and basic research (e.g., Bech, 2002; Johnson et al., 2008; Malkoff-Schwartz et al., 1998). Scores on the MAS reliably differentiate placebo and treatment groups, as well as detect changes in symptoms associated with treatment (Bech, 2002).

The SADS-C (Spitzer & Endicott, 1978) mania subscale is a five-item interview that assesses current severity of manic symptoms. Items are rated on a six-point scale that includes behavioral anchors. Good interrater reliability has been established in a range of settings with the exception of a sample of patients referred for emergency evaluation (intraclass correlation = .63 for mania; Rogers, Jackson, Salekin, & Neumann, 2003). Expected elevations on the scale have been seen in a bipolar sample compared to patients with other psychiatric disorders, as have robust correlations with another interview to assess manic severity, the MAS ($r = .89$; Johnson, Magaro, & Stern, 1986). Support for the scale in factor analytic studies has been mixed. One study found that all items loaded onto a single factor distinct from dysphoria, insomnia, and psychosis (Rogers et al., 2003). However, less factor analytic support was

obtained in a study that examined the item loadings for the SADS-C and a nurse observation scale for mania (Swann et al., 2001).

Self Report Measures

Two self-report measures of symptom severity have strong psychometric support: the Altman Self-Rating Mania (ASRM) Scale and the Self-Rating Mania Inventory (SRMI). We will also discuss other measures under development.

The ASRM scale (Altman, Hedeker, Peterson, & Davis, 1997) is a five-item scale that assesses mood, self-confidence, sleep disturbance, speech, and activity level over the past week. Items are scored on a 0 (absent) to 4 (present nearly all the time) scale, with total scores ranging from 0 to 20. Although the brevity can be an advantage, the scale covers fewer symptoms than other mania scales. Normative data for the ASRM have been gathered across major diagnostic groups (Altman et al., 1997; Altman, Hedeker, Peterson, & Davis, 2001).

The ASRM has demonstrated good psychometric properties. A cutoff score of 5.5 is recommended, as it has shown an optimal combination of sensitivity and specificity (85% and 86%, respectively). The ASRM also shows good sensitivity to treatment, with an average decrease of five points after discharge from the hospital (Altman et al., 2001). Finally, the ASRM demonstrated adequate internal consistency and concurrent validity when compared to SADS-based diagnoses, the YMRS (Young et al., 1978), and the Clinician-Administered Rating Scale for Mania (Altman et al., 1994, 1997, 2001). It should be noted that both of the published validation studies for the ASRM were conducted by the same research group. On the other hand, the scale has been shown to demonstrate expected correlations with psychological constructs related to mania, such as poor regulation of positive emotions (Feldman, Joormann, & Johnson, 2008).

The Self-Report Manic Inventory (SRMI; Braunig, Shugar, & Kruger, 1996; Shugar, Schertzer, Toner, & Di Gasbarro, 1992) is a 47-item true-false inventory that assesses increased energy, increased spending, increased sexual drive, increased verbosity, elation, irritability, racing thoughts and decreased concentration, grandiosity, and paranoid or psychotic experiences during the past week, and includes an item that addresses insight. Normative data have been reported in three small studies of inpatients, and these studies each provided estimates of good internal consistency (Altman et al., 2001; Braunig et al., 1996; Shugar et al., 1992). In two studies, the SRMI was found to have good discriminant validity, differentiating people with bipolar disorder from those with other psychopathology (Braunig et al., 1996; Shugar et al., 1992). However, another study found the SRMI to have low concurrent validity as compared to the ASRM (Altman et al., 2001). The scale appears sensitive to change in symptoms. It may not be well suited for inpatient assessment, however, because seven of the SRMI items describe behaviors that would not be possible within a hospital setting (Altman et al., 2001).

The Internal State Scale (ISS; Bauer et al., 1991) is a 17-item scale that discriminates mood state and tracks manic and depressive symptoms. There are four empirically derived subscales: Activation, Well-Being, Perceived Conflict, and Depression Index. The Activation subscale (five items) assesses racing thoughts and behavioral activation, specifically feeling restless, sped-up, overactive, and impulsive. These items appear to capture general arousal more than symptoms of mania. Still, the Activation correlates well with other measures of mania (Bauer, Vojsa, Kinoshian, Altshuler, & Glick, 2000). The overall scale has demonstrated correlations with other measures of mania ranging from .21 to .60 and rates of correct classification ranging from .55 to .78 (Altman et al., 2001; Bauer et al., 1991; Bauer et al., 2000; Cooke, Krüger, & Shugar, 1996). The measure is sensitive to symptom decreases during treatment (Altman et al., 2001; Bauer et al., 1991; Cooke et al., 1996). Despite these strengths, the ISS scale has a low sensitivity to manic symptoms at the time of hospitalization (Altman et al., 2001). In

addition, scoring algorithms vary substantially across studies, as do means and standard deviations of score distributions (Altman et al., 2001; Bauer et al., 1991; Cooke et al., 1996). Thus, the ISS is not currently recommended.

Self-Monitoring Tools

Continuous monitoring of symptoms and functioning is pivotal for people suffering from chronic, recurrent conditions like bipolar disorder (e.g., Horn et al., 2002; Schärer, Hartweg, Hoern, et al., 2002). Such frequent monitoring, however, can be expensive both economically and in terms of clinicians' time. In addition, there is increasing consensus regarding the benefits of a collaborative care model for bipolar disorder, in which patients play an active role in managing their illness (Bauer et al., 2006a, 2006b; Sajatovic et al., 2005). Enlisting patients' input can have numerous benefits, including reduced costs, higher patient investment in treatment, and higher validity than clinician observations alone. These benefits may be especially relevant for longitudinal data with high variability, such as may be seen with rapid-cycling patients (Lam & Wong, 2005). In addition to the tracking of bipolar symptoms such as sleep disturbance and mood, self-monitoring may also provide broader information regarding important issues such as medication adherence and psychosocial functioning. These facts have led to a growing literature supporting the use of self-monitoring tools for bipolar disorder. For instance, the NIMH prospective Life-Chart Method (NIMH-LCM-p) can provide detailed information regarding rapid fluctuations in mood (Denicoff et al., 2000, 2002).

Frequent monitoring of bipolar symptoms can produce so much data that entering and organizing it into a useful format may be incredibly time-consuming. In response to this, some research has focused on the use of palmtop computers and other electronic formats for self-monitoring. Examples include a palmtop version of the NIMH-LCM (Schärer, Hartweg, Valerius, et al., 2002) as well as ChronoRecord software, the latter of which has shown significant correlations with the YMRS (Bauer et al., 2008).

Most of the research in support of self-monitoring in bipolar disorder should be considered preliminary, but promising. In addition to the methods described above, many clients find it helpful to create their own self-monitoring forms or to complete brief checklists to track their progress over time. Many consumer-oriented websites, such as that maintained by the Depression and Bipolar Support Alliance, provide such forms. To increase awareness of symptoms, these self-monitoring forms can be compared to clinician-rated interviews. This is an important area for future study, and it is the hope of the authors that self-monitoring methods continue to be refined and validated for bipolar disorder.

Summary of Symptom Severity Measures

At least two interview measures (the YMRS and MAS), as well as some self-report measures (e.g., the Altman and SRMI), have received psychometric support. Self-report measures can be completed quickly, but brevity and ease of use may also result in reduced precision. Self-monitoring may also be useful to help increase awareness about symptoms and to track progress over time, but further research is required in this domain.

CONCLUSIONS AND FUTURE DIRECTIONS

This article has summarized assessment tools for screening, diagnosis, and symptom monitoring within bipolar disorder. We would note that there are many important aspects of assessment in bipolar disorder that we have not addressed. Although the symptom severity and diagnostic scales covered above predominately address manic symptoms, we urge readers to evaluate a broader range of outcomes, including depression, quality of life, and social functioning. People with bipolar disorder experience at least some depressive symptoms at

least one-third of the weeks in a year (Judd et al., 2002; Keck & McElroy, 2003), and these subsyndromal depressive symptoms can be associated with substantial impairment across a variety of domains (Altshuler et al., 2006). High risk for suicide has been documented during depression within bipolar disorder (Angst et al., 2005); thus it will also be important to assess for depressive symptoms and suicidality. To date, there is strong evidence that bipolar and unipolar depressive symptoms are relatively similar (Johnson & Kizer, 2002), so applying the well-validated measures of depression from the unipolar literature is a reasonable strategy. Patients report that improvement in quality of life is a more important treatment goal to them than are specific symptoms, highlighting the importance of this oft-ignored domain (Michalak, Yatham, Kolesar, & Lam, 2006). Whereas measures of these constructs have been developed for other disorders such as depression and schizophrenia, this is a realm that remains largely untapped for bipolar disorder, with at least one exception (e.g., Michalak et al., 2006). In addition, there is some debate regarding the ultimate treatment goals for bipolar disorder. Given the high base rates of subsyndromal symptoms, complete recovery may be an unrealistic goal, or require levels of medication that would lead to intolerable side effects (Sachs & Rush, 2003). Proper care must take individual needs into account, but to date little research has directly addressed this issue. Overall, it is highly recommended that researchers and clinicians pay attention to issues that extend far beyond the level of mania. For those who seek a more detailed review of assessment measures for bipolar disorder or psychiatric conditions more generally, the authors recommend comprehensive books such as the *Handbook of Psychiatric Measures* (Rush, First, & Blacker, 2008).

Returning to the focus of this article, though, the good news is that well-validated tools exist for the assessment of mania in adults. Reliable and valid measures are available for the diagnosis of bipolar I disorder, and indeed, the psychometric characteristics of these tools are as good as those seen for most Axis I disorders. Similarly, scales are available to measure symptoms using both interviewer and client perspectives.

On the other hand, much work remains to be done in this domain. A first goal would be the refinement of diagnostic measures for bipolar II disorder and other milder forms of bipolar disorder. Ideally, research and dialogue in the near future will help to establish accepted standards for defining hypomanic episodes. A second major goal is the refinement of screening tools. With the possible exception of the GBI, no self-report measure has consistently achieved acceptable levels of sensitivity and specificity within community samples, and conclusions regarding the GBI are limited by the existence of several different versions and cutoffs. One might expect that the most pressing need would be for screening tools that were viable for community or outpatient screening, as by the time a person is hospitalized, symptoms may be so extreme as to be easily diagnosed. A third major goal is more systematic research on how to integrate clinician and self-report ratings of symptom severity, especially in the face of potentially impaired insight for those with bipolar disorder (Ghaemi, Boiman, & Goodwin, 2000). Intriguingly, although researchers have now begun to examine the relative weight to give ratings from different informants in understanding juvenile bipolar disorder (Findling et al., 2002), such research has not been conducted in adult bipolar disorder. Rather, researchers focused on adult bipolar disorder have often failed to take into account patient perspectives on severity. We are hopeful that future research will continue to refine this field, and that this review has illuminated research challenges to be tackled.

References

- Akiskal, HS. Classification, diagnosis, and boundaries of bipolar disorders: A review. In: Maj, M.; Akiskal, HS.; Lopez-Ibor, JJ.; Sarotius, N., editors. *Bipolar disorder*. Chichester, UK: Wiley; 2002. p. 1-52.

- Akiskal HS, Akiskal KK. TEMPS: Temperament evaluation of Memphis, Pisa, Paris and San Diego. *Journal of Affective Disorders* 2005;85:1–2. [PubMed: 15780670]
- Akiskal HS, Benazzi F. Optimizing the detection of bipolar II disorder in outpatient private practice: Toward a systematization of clinical diagnostic wisdom. *Journal of Clinical Psychiatry* 2005;66:914–921. [PubMed: 16013908]
- Akiskal HS, Maser JD, Zeller PJ, Endicott J, Coryell W, Keller M, et al. Switching from “unipolar” to bipolar II. An 11-year prospective study of clinical and temperamental predictors in 559 patients. *Archives of General Psychiatry* 1995;52:114–123. [PubMed: 7848047]
- Akiskal HS, Mendlowicz MV, Jean-Louis G, Rapaport MH, Kelsoe JR, Gillin JC, Smith TL. TEMPS-A: Validation of a short version of a self-rated instrument designed to measure variations in temperament. *Journal of Affective Disorders* 2005;85:45–52. [PubMed: 15780675]
- Altman EG, Hedeker DR, Janicak P, Peterson JL, Davis JM. The Clinician-Administered Rating Scale for Mania (CARS-M): Development, reliability, and validity. *Biological Psychiatry* 1994;36:124–134. [PubMed: 7948445]
- Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman Self-Rating Mania Scale. *Biological Psychiatry* 1997;42:948–955. [PubMed: 9359982]
- Altman EG, Hedeker D, Peterson JL, Davis JM. A comparative evaluation of three self-rating scales for acute mania. *Biological Psychiatry* 2001;50:468–471. [PubMed: 11566165]
- Altshuler LL, Post RM, Black DO, Keck PE, Nolen WA, Frye MA, et al. Subsyndromal depressive symptoms are associated with functional impairment in patients with bipolar disorder: Results of a large multisite study. *Journal of Clinical Psychiatry* 2006;67:1551–1560. [PubMed: 17107246]
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: Author; 2000. text rev
- Andreasen NC, Grove WM, Shapiro RW, Keller MB, Hirschfeld RM, McDonald-Scott P. Reliability of lifetime diagnosis. A multicenter collaborative perspective. *Archives of General Psychiatry* 1981;38:400–405. [PubMed: 7212970]
- Angst J, Adolfsson R, Benazzi F, Gamma A, Hantouche E, Meyer TD, et al. The HCL-32: Towards a self-assessment tool for hypomanic symptoms in outpatients. *Journal of Affective Disorders* 2005;88:217–233. [PubMed: 16125784]
- Basco MR, Bostie JQ, Davies D, Rush AJ, Witte B, Hendrickse W, et al. Methods to improve diagnostic accuracy in a community mental health setting. *American Journal of Psychiatry* 2000;157:1599–1605. [PubMed: 11007713]
- Bauer MS, Crits-Christoph P, Ball WA, Dewees E, McAllister T, Alahi P, et al. Independent assessment of manic and depressive symptoms by self-rating: Scale characteristics and implications for the study of mania. *Archives of General Psychiatry* 1991;48:807–812. [PubMed: 1929771]
- Bauer MS, McBride L, Williford WO, Glick H, Kinosian B, Altshuler L, et al. Collaborative care for bipolar disorder: Part I: Intervention and implementation in a randomized effectiveness trial. *Psychiatric Services* 2006a;57:927–936. [PubMed: 16816276]
- Bauer MS, McBride L, Williford WO, Glick H, Kinosian B, Altshuler L, et al. Collaborative care for bipolar disorder: Part II: Impact on clinical outcome, function, and costs. *Psychiatric Services* 2006b; 57:937–945. [PubMed: 16816277]
- Bauer MS, Vojta C, Kinosian B, Altshuler L, Glick H. The Internal State Scale: Replication of its discriminating abilities in a multisite, public sector sample. *Bipolar Disorders* 2000;2:340–346. [PubMed: 11252648]
- Bauer MS, Wilson T, Neuhaus K, Sasse J, Pfennig A, Lewitzka U, Grof P, et al. Self-reporting software for bipolar disorder: Validation of ChronoRecord by patients with mania. *Psychiatry Research* 2008;159:359–366. [PubMed: 18423616]
- Bech P. Rating scales for mood disorders: Applicability, consistency and construct validity. *Acta Psychiatrica Scandinavica* 1988;78:45–55.
- Bech P. The Bech-Rafaelsen Mania Scale in clinical trials of therapies for bipolar disorder. *CNS Drugs* 2002;16:47–63. [PubMed: 11772118]
- Bech P, Bolwig TG, Kramp P, Rafaelsen OJ. The Bech-Rafaelsen Mania Scale and the Hamilton Depression Scale. *Acta Psychiatrica Scandinavica* 1979;59:420–430. [PubMed: 433633]

- Braünig P, Shugar G, Krüger S. An investigation of the Self-Report Mania Inventory as a diagnostic and severity scale for mania. *Comprehensive Psychiatry* 1996;37:52–55. [PubMed: 8770527]
- Brickman A, LoPiccolo C, Johnson SL. Screening for bipolar disorder by community providers [Letter to the editor]. *Psychiatric Services* 2002;53:349. [PubMed: 11875241]
- Carlson GA, Goodwin FK. The stages of mania: A longitudinal analysis of the manic episode. *Archives of General Psychiatry* 1973;28:221–228. [PubMed: 4684288]
- Cooke RG, Krüger S, Shugar G. Comparative evaluation of two self-report mania rating scales. *Biological Psychiatry* 1996;40:279–283. [PubMed: 8871774]
- Coryell W, Endicott J, Maser JD, Keller MB, Leon AC, Akiskal HS. Long-term stability of polarity distinctions in the affective disorders. *The American Journal of Psychiatry* 1995;152:385–390. [PubMed: 7864264]
- Das AK, Olfson M, Gameroff MJ, Pilowsky DJ, Blanco C, Feder A, et al. Screening for bipolar disorder in a primary care practice. *Journal of the American Medical Association* 2005;293:956–963. [PubMed: 15728166]
- Dell’Osso L, Armani A, Rucci P, Frank E, Fagiolini A, Corretti G, et al. Measuring mood spectrum: Comparison of interview (SCI-MOODS) and self-report (MOODS-SR) instruments. *Comprehensive Psychiatry* 2002;43:69–73. [PubMed: 11788923]
- Denicoff KD, Ali SO, Sollinger AB, Smith-Jackson EE, Leverich GS, Post RM. Utility of the daily prospective National Institute of Mental Health Life-Chart Method (NIMH-LCM-p) ratings in clinical trials of bipolar disorder. *Depression and Anxiety* 2002;15:1–9. [PubMed: 11816046]
- Denicoff KD, Leverich GS, Nolen WA, Rush AJ, McElroy SL, Keck PE Jr, et al. Validation of the prospective NIMH Life-Chart Method (NIMH-LCM-p) for longitudinal assessment of bipolar illness. *Psychological Medicine* 2000;30:1391–1397. [PubMed: 11097079]
- Depue, RA.; Klein, DN. *Relatives at risk for mental disorder*. New York: Raven Press; 1988. Identification of unipolar and bipolar affective conditions in nonclinical and clinical populations by the General Behavior Inventory; p. 179-204.
- Depue RA, Krauss S, Spoont MR, Arbisi P. General Behavior Inventory identification of unipolar and bipolar affective conditions in a nonclinical university population. *Journal of Abnormal Psychology* 1989;98:117–126. [PubMed: 2708652]
- Depue RA, Slater JF, Wolfstetter-Kausch H, Klein D, Goplerud E, Farr D. A behavioral paradigm for identifying persons at risk for bipolar depressive disorder: A conceptual framework and five validation studies. *Journal of Abnormal Psychology* 1981;90:381–437. [PubMed: 7298991]
- Double DB. The factor structure of mania rating scales. *Journal of Affective Disorders* 1990;18:113–119. [PubMed: 2137468]
- Dunner, DL. Bipolar depression with hypomania (bipolar II) DSM-IV sourcebook. Widiger, TA.; Frances, AJ.; Pincus, HA.; Ross, R.; First, MB.; Davis, WW., editors. Vol. 2. Washington, DC: American Psychiatric Association Press; 1996. p. 53-64.
- Dunner DL, Tay LK. Diagnostic reliability of the history of hypomania in bipolar II patients and patients with major depression. *Journal of Comprehensive Psychiatry* 1993;34:303–307.
- Eckblad M, Chapman LJ. Development and validation of a scale for hypomanic personality. *Journal of Abnormal Psychology* 1986;95:214–222. [PubMed: 3745642]
- Endicott J, Spitzer RL. A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry* 1978;35:837–844. [PubMed: 678037]
- Feldman GC, Joermann J, Johnson SL. Responses to positive affect: A self-report measure of rumination and dampening. *Cognitive Therapy and Research* 2008;32:507–525. [PubMed: 20360998]
- Findling RL, Youngstrom EA, Danielson CK, DelPorto-Bedoya D, Papish-David R, Townsend L, et al. Clinical decision-making using the General Behavior Inventory in juvenile bipolarity. *Bipolar Disorders* 2002;4:34–42. [PubMed: 12047493]
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. *Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I), Clinician Version*. Washington, DC: American Psychiatric Association; 1997.
- Ghaemi SN, Boiman E, Goodwin FK. Insight and outcome in bipolar, unipolar, and anxiety disorders. *Comprehensive Psychiatry* 2000;41:167–171. [PubMed: 10834624]
- Ghaemi SN, Lenox MS, Baldessarini RJ. Effectiveness and safety of long-term antidepressant treatment in bipolar disorder. *Journal of Clinical Psychiatry* 2001;62:565–569. [PubMed: 11488370]

- Ghaemi SN, Miller CJ, Berv DA, Klugman J, Rosenquist KJ, Pies RW. Sensitivity and specificity of a new bipolar spectrum diagnostic scale. *Journal of Affective Disorders* 2005;84:273–277. [PubMed: 15708426]
- Hantouche EG, Angst J, Lancrenon S, Gerard D, Allilaire JF. Feasibility of auto-evaluation in the detection of hypomania. *Annales Medico-Psychologiques* 2006;164:721–725.
- Hirschfeld RMA, Holzer C, Calabrese JR, Weissman M, Reed M, Davies M, et al. Validity of the Mood Disorder Questionnaire: A general population study. *American Journal of Psychiatry* 2003;160:178–180. [PubMed: 12505821]
- Hirschfeld RMA, Williams JBW, Spitzer RL, Calabrese JR, Flynn L, Keck PE Jr, et al. Development and validation of a screening instrument for bipolar spectrum disorder: The Mood Disorder Questionnaire. *American Journal of Psychiatry* 2000;157:1873–1875. [PubMed: 11058490]
- Hörn M, Schärer L, Walser S, Scherer-Klabunde D, Biedermann C, Walden J. Comparison of long-term monitoring methods for bipolar affective disorder. *Neuropsychobiology* 2002;45(Suppl 1):27–32.
- Isometsä E, Suominen K, Mantere O, Valtonen H, Leppämäki S, Pippingsköld M, et al. The Mood Disorder Questionnaire improves recognition of bipolar disorder in psychiatric care. *BMC Psychiatry* 2003;3. [PubMed: 12659652]
- Johnson SL. Mania and dysregulation in goal pursuit. *Clinical Psychology Review* 2005;25:241–262. [PubMed: 15642648]
- Johnson SL, Cuellar AK, Ruggero C, Winett-Perlman C, Goodnick P, White R, et al. Life events as predictors of mania and depression in bipolar I disorder. *Journal of Abnormal Psychology* 2008;117:268–277. [PubMed: 18489203]
- Johnson, SL.; Kizer, A. Bipolar and unipolar depression: A comparison of clinical phenomenology and psychosocial predictors. In: Gotlib, IH.; Hammen, CL., editors. *Handbook of depression*. New York: Guilford Press; 2002. p. 141-165.
- Johnson, SL.; Leahy, RL., editors. *Psychological treatment of bipolar disorder*. New York: Guilford Press; 2004.
- Johnson MH, Magaro PA, Stern SL. Use of the SADS-C as a diagnostic and symptom severity measure. *Journal of Consulting and Clinical Psychology* 1986;54:546–551. [PubMed: 3745610]
- Johnson, SL.; Miller, CJ.; Eisner, LR. Bipolar disorder. In: Hunsley, J.; Mash, EJ., editors. *Guide to assessments that work*. New York: Oxford University Press; 2008.
- Judd LL, Akiskal HS, Schettler PJ, Endicott J, Maser J, Solomon DA, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Archives of General Psychiatry* 2002;59:530–537. [PubMed: 12044195]
- Karam EG, Mneimneh ZN, Salamoun MM, Akiskal HS, Akiskal KK. Suitability of the TEMPS-A for population-based studies: Ease of administration and stability of affective temperament in its Lebanese version. *Journal of Affective Disorders* 2007;98:45–53. [PubMed: 16949160]
- Karkowski LM, Kendler KS. An examination of the genetic relationship between bipolar and unipolar illness in an epidemiological sample. *Psychiatric Genetics* 1997;7:159–163. [PubMed: 9460800]
- Keck PE, McElroy SL. New approaches in managing bipolar depression. *Journal of Clinical Psychiatry* 2003;64:13–18. [PubMed: 12625800]
- Keller MB, Lavori PW, McDonald-Scott P, Scheftner WA, Andreasen NC, Shapiro RW, et al. Reliability of lifetime diagnoses and symptoms in patients with current psychiatric disorder. *Journal of Psychiatry Research* 1981;16:229–240.
- Kesebir S, Vahip S, Akdeniz F, Yuncu Z, Alkan M, Akiskal HS. Affective temperaments as measured by TEMPS-A in patients with bipolar I disorder and their first-degree relatives: A controlled study. *Journal of Affective Disorders* 2005;85:127–133. [PubMed: 15780683]
- Kessler RC, Akiskal HS, Angst J, Guyer M, Hirschfeld RMA, Merikangas KR, et al. Validity of the assessment of bipolar spectrum disorders in the WHO CIDI 3.0. *Journal of Affective Disorders* 2006;96:259–269. [PubMed: 16997383]
- Kessler RC, Berglund P, Demler O, Jin R, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey replication. *Archives of General Psychiatry* 2005;62:593–602. [PubMed: 15939837]

- Kessler RC, Rubinow DR, Holmes C, Abelson JM, Zhao S. The epidemiology of DSM-III-R bipolar I disorder in a general population survey. *Psychological Medicine* 1997;27:1079–1089. [PubMed: 9300513]
- Kessler, RC.; Zhao, S. The prevalence of mental illness. In: Horowitz, AV.; Scheid, TL., editors. *A handbook for the study of mental health: Social contexts, theories, and systems*. Cambridge, UK: Cambridge University Press; 1999. p. 58-78.
- Kieseppä T, Partonen T, Haukka J, Kaprio J, Lönqvist J. High concordance of bipolar I disorder in a nationwide sample of twins. *American Journal of Psychiatry* 2004;161:1814–1821. [PubMed: 15465978]
- Klein DN, Dickstein S, Taylor EB, Harding K. Identifying chronic affective disorders in out-patients: Validation of the general behavior inventory. *Journal of Consulting and Clinical Psychology* 1989;57:106–111. [PubMed: 2925959]
- Kochman FJ, Hantouche EG, Ferrari P, Lancrenon S, Bayart D, Akiskal HS. Cyclothymic disorder temperament as a prospective predictor of bipolarity and suicidality in children and adolescents with major depressive disorder. *Journal of Affective Disorders* 2005;85:181–189. [PubMed: 15780688]
- Kraemer, HC. *Evaluating medical tests*. Newbury Park, CA: Sage; 1992.
- Kwapil, TR. Performance of the Hypomanic Personality Scale in a clinical sample. 2008. Unpublished raw data
- Kwapil TR, Miller MB, Zinser MC, Chapman LJ, Chapman J, Eckblad M. A longitudinal study of high scorers on the hypomanic personality scale. *Journal of Abnormal Psychology* 2000;109:222–226. [PubMed: 10895560]
- Lam D, Wong. Prodromes, coping strategies, and psychological interventions in bipolar disorders. *Clinical Psychology Review* 2005;25:1028–1042. [PubMed: 16125292]
- Leibenluft E, Albert PS, Rosenthal NE, Wehr TA. Relationship between sleep and mood in patients with rapid-cycling bipolar disorder. *Psychiatry Research* 1996;63:161–168. [PubMed: 8878312]
- Licht RW, Jensen J. Validation of the Bech-Rafaelsen Mania Scale using latent structure analysis. *Acta Psychiatrica Scandinavica* 1997;96:367–372. [PubMed: 9395155]
- Lish JD, Dime-Meenan S, Whybrow PC, Price RA, Hirschfeld RM. The National Depressive and Manic-Depressive Association (DMDA) survey of bipolar members. *Journal of Affective Disorders* 1994;31:281–294. [PubMed: 7989643]
- Malkoff-Schwartz S, Frank E, Anderson B, Sherrill JT, Siegel L, Patterson D, et al. Stressful life events and social rhythm disruption in the onset of manic and depressive bipolar episodes. *Archives of General Psychiatry* 1998;55:702–707. [PubMed: 9707380]
- Mallon JC, Klein DN, Bornstein RF, Slater JF. Discriminant validity of the General Behavior Inventory: An outpatient study. *Journal of Personality Assessment* 1986;50:568–577. [PubMed: 3820050]
- Mantere O, Suominen K, Leppamäki S, Arvilommi P, Isometsä E. The clinical characteristics of DSM-IV bipolar I and II disorders: Baseline findings from the Jorvi Bipolar Study (JoBS). *Bipolar Disorders* 2004;6:395–405. [PubMed: 15383132]
- Matsumoto S, Akiyama T, Tsuda H, Miyake Y, Kawamura Y, Noda T, et al. Reliability and validity of TEMPS-A in a Japanese non-clinical population: Application to unipolar and bipolar depressives. *Journal of Affective Disorders* 2005;85:85–92. [PubMed: 15780679]
- Mendlowicz MV, Jean-Louis G, Kelsoe JR, Akiskal HS. A comparison of recovered bipolar patients, healthy relatives of bipolar probands, and normal controls using the short TEMPS-A. *Journal of Affective Disorders* 2005;85:147–151. [PubMed: 15780685]
- Meyer TD, Hammelstein P, Nilsson LG, Skeppar P, Adolfsson R, Angst J. The Hypomania Checklist (HCL-32): Its factorial structure and association to indices of impairment in German and Swedish nonclinical samples. *Comprehensive Psychiatry* 2007;48:79–87. [PubMed: 17145286]
- Michalak E, Yatham L, Kolesar S, Lam R. Bipolar disorder and quality of life: A patient-centered perspective. *Quality of Life Research* 2006;15:25–37. [PubMed: 16411028]
- Miller, CJ.; Johnson, SL.; Carver, CS. Unpublished manuscript. 2008. Testing the utility of self-report screens to detect bipolar disorder among undergraduates.
- Miller CJ, Klugman J, Berv DA, Rosenquist KJ, Ghaemi SN. Sensitivity and specificity of the Mood Disorder Questionnaire for detecting bipolar disorder. *Journal of Affective Disorders* 2004;81:167–171. [PubMed: 15306144]

- Phelps JR, Ghaemi SN. Improving the diagnosis of bipolar disorder: Predictive value of screening tests. *Journal of Affective Disorders* 2006;92:141–148. [PubMed: 16529822]
- Prien RF, Potter WZ. NIMH workshop report on treatment of bipolar disorder. *Psychopharmacology Bulletin* 1990;26:409–427. [PubMed: 2087538]
- Reigier DA, Farmer ME, Rae DS, Locke BZ, Keith SJ, Judd LL, et al. Comorbidity of mental disorders with alcohol and other substance abuse: Results from the Epidemiological Catchment Area (ECA) Study. *Journal of the American Medical Association* 1990;264:2511–2518. [PubMed: 2232018]
- Rice JP, McDonald-Scott P, Endicott J, Coryell W, Grove WM, Keller MB, et al. The stability of diagnosis with an application to bipolar II disorder. *Journal of Psychiatry Research* 1986;19:285–296.
- Rizvi S, Zaretsky AE. Psychotherapy through the phases of bipolar disorder: Evidence for general efficacy and differential effects. *Journal of Clinical Psychology* 2007;63:491–506. [PubMed: 17417815]
- Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, Farmer A, et al. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry* 1988;45:1069–1077. [PubMed: 2848472]
- Rogers, R.; Jackson, R.L.; Cashel, M. The Schedule for Affective Disorders and Schizophrenia (SADS). In: Rogers, R., editor. *Handbook of diagnostic and structural interviewing*. New York: Guilford Press; 2001. p. 84-102.
- Rogers R, Jackson RL, Salekin KL, Neumann CS. Assessing axis I symptomatology on the SADS-C in two correctional samples: The validation of subscales and a screen for malingered presentations. *Journal of Personality Assessment* 2003;81:281–290. [PubMed: 14638453]
- Rush, AJ., Jr; First, MB.; Blacker, D. *Handbook of psychiatric measures*. 2. Washington, DC: American Psychiatric Publishing; 2008.
- Sachs GS, Rush AJ. Response, remission, and recovery in bipolar disorders: What are the realistic treatment goals? *Journal of Clinical Psychiatry* 2003;64(Suppl 6):18–22. [PubMed: 12720476]
- Sachs GS, Thase ME, Otto MW, Bauer M, Miklowitz D, Wisniewski SR, et al. Rationale, design, and methods of the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Biological Psychiatry* 2003;53:1028–1042. [PubMed: 12788248]
- Sajatovic M, Davies M, Bauer MS, McBride L, Hays RW, Safavi R, et al. Attitudes regarding the collaborative practice model and treatment adherence among individuals with bipolar disorder. *Comprehensive Psychiatry* 2005;46:272–277. [PubMed: 16175758]
- Sandor R, Rihmer A, Ko N, Gonda X, Szili I, Szadoczky E, et al. Affective temperaments: Psychometric properties of the Hungarian TEMPS-A. *Psychiatria Hungarica* 2006;21:147–160.
- Schärer LO, Hartweg V, Hoern M, Graesslin Y, Strobl N, Frey S, et al. Electronic diary for bipolar patients. *Neuropsychobiology* 2002;46(Suppl 1):10–12. [PubMed: 12571426]
- Schärer LO, Hartweg V, Valerius G, Graf M, Hoern M, Biedermann C, et al. Life charts on a palmtop computer: First results of a feasibility study with an electronic diary for bipolar patients. *Bipolar Disorders* 2002;4(Suppl 1):107–108. [PubMed: 12479693]
- Secunda SK, Katz MM, Swann A, Koslow SH, Maas JW, Chuang S. Mania: Diagnosis, state measurement and prediction of treatment response. *Journal of Affective Disorders* 1985;8:113–121. [PubMed: 3157719]
- Shugar G, Schertzer S, Toner BB, Di Gasbarro I. Development, use, and factor analysis of a self-report inventory for mania. *Comprehensive Psychiatry* 1992;33:325–331. [PubMed: 1395552]
- Simpson SG, McMahon FJ, McInnis MG, MacKinnon DF, Edwin D, Folstein SE, et al. Diagnostic reliability of bipolar II disorder. *Archives of General Psychiatry* 2002;59:736–740. [PubMed: 12150650]
- Spitzer, RA.; Endicott, J. *Schedule for Affective Disorders and Schizophrenia—Change version*. New York: Biometrics Research, State Psychiatric Institute; 1978.
- Spitzer RL, Williams JBW, Gibbon M, First MB. The structured clinical interview for DSM-III-R (SCID). I. History, rationale, and description. *Archives of General Psychiatry* 1992;49:624–629. [PubMed: 1637252]
- Swann AC, Janicak PL, Calabrese JR, Bowden CL, Dilsaver SC, Morris DD, et al. Structure of mania: Depressive, irritable, and psychotic clusters with different retrospectively-assessed course patterns

of illness in randomized clinical trial participants. *Journal of Affective Disorders* 2001;67:123–132. [PubMed: 11869759]

Vazquez GH, Susana N, Mercado B, Romero E, Tifner S, Ramon MDL, et al. Validation of the TEMPS-A Buenos Aires: Spanish psychometric validation of affective temperaments in a population study of Argentina. *Journal of Affective Disorders* 2007;100:23–29. [PubMed: 17306377]

Vernon SW, Roberts RE. Use of the SADS-RDC in a tri-ethnic community survey. *Archives of General Psychiatry* 1982;39:47–52. [PubMed: 6976778]

Vieta E, Sánchez-Moreno J, Bulbena A, Chamorro L, Ramos JL, Artal J, et al. Cross validation with the Mood Disorder Questionnaire (MDQ) of an instrument for the detection of hypomania in Spanish: The 32 item hypomania symptom check list (HCL-32). *Journal of Affective Disorders* 2007;101:43–55. E. Vieta, and for the EDHIPO (hypomania detection study) group. [PubMed: 17189651]

Weber Rouget B, Gervasoni N, Dubuis V, Gex-Fabry M, Bondolfi G, Aubry J. Screening for bipolar disorders using the French version of the Mood Disorder Questionnaire (MDQ). *Journal of Affective Disorders* 2005;88:103–108. [PubMed: 16054228]

Williams JBW, Gibbon M, First MB, Spitzer RL, Davies M, Borus J, et al. The structured clinical interview for the DSM-III-R (SCID). II. Multisite test–retest reliability. *Archives of General Psychiatry* 1992;49:630–636. [PubMed: 1637253]

Winokur, G.; Clayton, PJ.; Reich, T. Manic depressive illness. St. Louis, MO: C. V. Mosby; 1969.

Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: Reliability, validity and sensitivity. *British Journal of Psychiatry* 1978;133:429–435. [PubMed: 728692]

Zimmerman M, Mattia JJ. Psychiatric diagnosis in clinical practice: Is comorbidity being missed? *Comprehensive Psychiatry* 1999;40:182–191. [PubMed: 10360612]

Table 1

Summary of Validated Bipolar Disorder Assessment Tools for Diagnosis

Name	Primary setting	Format	Notes
Structured Clinical Interview for <i>DSM-IV</i>	Both clinical and research settings	Semistructured interview	Consists of separate modules relevant to particular disorders; can be relatively time-consuming; good reliability/validity data
Schedule for Affective Disorders and Schizophrenia—Lifetime Version	Both clinical and research settings	Semistructured interview	Consists of separate modules relevant to particular disorders; can be relatively time-consuming; good reliability/validity data
General Behavior Inventory	Research settings	Self-report	Different versions have different lengths and cutoffs (e.g., 52- to 73-item length); items can be hard to understand due to complex wording and syntax; publicly available
Mood Disorder Questionnaire	Research settings	Self-report	Very brief inventory (15 items); appears more useful in clinical rather than community populations; publicly available

Table 2

Summary of Validated Bipolar Disorder Assessment Tools for Symptom Severity

Name	Primary setting	Format	Notes
Young Mania Rating Scale	Treatment trials	Clinician interview	Brief (15–30 minute) interview covers most, but not all, manic symptoms; publicly available
Bech-Rafaelsen Mania Rating Scale	Treatment trials	Clinician interview	Brief (10–20 minute) interview
Schedule for Affective Disorders and Schizophrenia—Change version mania subscale	Both clinical and research settings	Clinician interview	Very brief (five-item) interview with mixed factor analytic support
Altman Self-Rating Mania Scale	Research settings	Self-report	Five-item scale with good psychometric support, although items do not cover all manic symptoms; publicly available
Self-Report Manic Inventory	Research settings	Self-report	True–false format (48 items); not recommended for use with inpatients