A review of self-report and interview-based instruments to assess mania and hypomania symptoms

Una rassegna degli strumenti autovalutativi ed eterovalutativi per valutare i sintomi maniacali e ipomaniacali

P. Rucci, S. Calugi, M. Miniati[1], A. Fagiolini[2]

Department of Biomedical and Neuromotor Sciences, Alma Mater Studiorum University of Bologna; [1] Department of Psychiatry, Neurobiology, Pharmacology and Biotechnology, University of Pisa, Italy; [2] Department of Mental Health and Neuroscience, University of Siena School of Medicine

Summary

Objective

The aim of this paper is to provide an overview of the self-report and interview-based instruments to assess mania/hypomania symptoms and related features, with a focus on 7 selected instruments in widespread use to illustrate their psychometric properties, comparative performance and pros and cons.

Methods

A systematic search strategy was devised and queried on Medline from 1973 to 2012 using the terms mania, hypomania, instrument, scale, questionnaire, interview, validity, reliability, psychometric properties and adults, elderly, aged. To be included, a study had to be published in a peer-reviewed journal or book in English or Italian.

Results

Of the 17 self-report instruments identified, two (the Mood Disorder Questionnaire (MDQ) and the Hypomania Checklist-32 (HCL-32), received the most research attention. Although the psychometric properties of these instruments are good, their use as screening instruments to detect hypomania in the community or in patients with depression is partially limited by their low positive predictive value, related to the low prevalence of this condition. Nonetheless, they can be efficiently used to rule out the presence of hypomania. The Altman Self-Rating Mania Scale is increasingly being used to monitor mania symptoms over time by phone or email in patients diagnosed with bipolar disorder because it consists of only 5 items. When the aim is early detection of manic/hypomanic symptoms that a patient may have experienced during their lifetime, the 33-item subset of the MOODS-SR seems promising because it includes the key psychopathology dimensions that better discriminate bipolar from unipolar disorder.

Of the interview-based instruments, the Young Mania Rating Scale and the Bech-Rafaelsen Mania Scale are the most widely used outcome measures in clinical trials. Although they were developed more than 30 years ago, they continue to be the gold standard for research purposes. The two instruments have a similar coverage, although the YMRS is preferred over the BR-MAS because it includes an item on insight.

Conclusions

Although no instrument can replace the need for accurate clinical diagnosis based on patient history, we argue that the increasing use of self-report instruments to screen bipolar disorder in patients presenting with depression or to monitor mania/hypomania symptoms over time may contribute to increasing the use of routine standardized assessment. Measurement-based care as the standard of care has the potential to transform psychiatric practice, move psychiatry into the mainstream of medicine, and ultimately improve the quality of care for patients with psychiatric illness.

Key words

Mania • Hypomania • Bipolar spectrum • Rating scales • Interview • Questionnaire • Validity • Reliability

Introduction

Bipolar disorder is a serious illness associated with significant psychosocial morbidity and excess mortality. Recent research carried out by World Health Organization World Mental Health Survey Initiative in community adults from 11 countries worldwide indicated that bipolar disorder, when defined to include milder variants such as bipolar II disorder and subthreshold bipolar disorder, has a lifetime prevalence of 2.4% [1]. Studies carried out in psychiatric and primary care settings have found that bipolar disorder is sometimes under-recognized, particularly in patients presenting for treatment of depression [2-5]. Even for those patients diagnosed with bipolar disorder, the time lag between initial treatment seeking and correct diagnosis often exceeds 10 years [6][7].

The treatment and clinical implications of the failure to recognize bipolar disorder in depressed patients include the under-prescription of mood stabilizers, an increased

Correspondence

Paola Rucci, Department of Biomedical and Neuromotor Sciences, Alma Mater Studiorum, University of Bologna, via San Giacomo 12, 40126 Bologna, Italy • Tel. +39 051 2094837 • Fax 051-2094839 • E-mail: paola.rucci2@unibo.it

risk of rapid, cycling and increased costs of care [7-10]. When symptomatic, patients with bipolar disorder are much more likely to experience symptoms of depression and anxiety rather than symptoms of mania or hypomania [11]. It is therefore frequent that, when presenting for treatment, patients with bipolar disorder are not in the manic or hypomanic phases of the illness. This suggests that manic phases, especially when brief or not characterized by impulse dyscontrol, need to be elicited with retrospective assessment, considering the frequent lack of subjective suffering, enhanced productivity, ego-syntonicity and diurnal or seasonal rhythmicity associated with several manic/hypomanic symptoms [12].

Recommendations for improving the detection of bipolar disorder include careful clinical evaluations inquiring about a history of mania and hypomania and the use of screening questionnaires [13-16]. A systematic classification of self-report and interview-based instruments to assess mania and hypomania might help the clinician to make the most appropriate instrument selection for the different research and clinical purposes [17].

In 2009, Picardi [18] reviewed the rating scales for bipolar disorder, according to the type of symptoms to be assessed (depressive, manic, psychotic) and the purpose of the instrument (screening, early identification). The aim of this paper is to provide a broader overview of the existing self-report questionnaires and of interview-based clinical instruments to assess mania symptoms and related features, with a focus on 7 selected instruments in widespread use to illustrate their comparative performance and pros and cons.

Methods

A systematic search strategy was devised and queried on Medline from 1973 to 2012 including the terms mania, hypomania, bipolar spectrum, mood spectrum, instrument, rating scale, questionnaire, interview and validity, reliability, psychometric properties.

To be included, a study had to be published in a peer-reviewed journal or book and in English.

Results

The Medline search yielded a total of 43 studies, retrieved from journal articles, describing 31 instruments, 17 self-report and 14 interview-based.

Table I summarizes the characteristics and the psychometric properties of the instruments identified, including the internal consistency, concurrent/discriminant validity, inter-rater reliability and factor structure (when applicable and when available). The assessment instruments for manic symptoms are classified according to their for-

mat (self-report or interview-based) and are sorted in decreasing order by year of publication. The seven selected instruments are in boldface.

Self-report questionnaires

The first author who strongly supported the use of self-report rating scales to assess the presence and/or severity of manic symptoms was Altman, who in 1997 developed the Altman Self-Rating Mania Scale [19], consisting of 5 items rated on a Likert scale of 0-4. In his commentary in 1998 [20], he examined the extent to which the severity of illness, the presence of psychosis or the lack of insight may threaten the reliability of a self-report measure. After comparing his measure against interview-based instruments (CARS-M, MRS), he concluded that 'self-rating mania scales are both reliable and valid for patients with manic symptoms, including those with psychotic features and those having little or no insight into their illness.'

This scale has recently gained a renewed popularity and is used in the US, together with the Quick Inventory of Depressive Symptoms (QIDS), to monitor manic and depressive symptoms over time through weekly text messages and e-mails [21][22]. These studies suggested that text message-based symptom monitoring during routine follow-up may be a reliable alternative to in-person interviews.

Among the self-report instruments developed to improve the detection of bipolar disorders, the Mood Disorder Questionnaire (MDQ) and the Hypomania Checklist-32 (HCL-32) have currently received most research attention.

Below we summarize the characteristics and psychometric properties of these 2 scales and the Bipolar Spectrum Diagnostic Scale (BSDS) and the Mood Spectrum Questionnaire (MOODS-SR), which were developed to detect softer forms of bipolar disorders or isolated symptoms of manic-hypomanic spectrum co-occurring with other axis-I disorders.

The other instruments listed in Table I include the Self-Report Manic Inventory (SRMI) and the Altman Self-rating Mania Scale (ASRM) designed specifically for assessing the manic pole of the illness and seven bipolarity rating scales: Visual Analogue Mood Scale (VAMS), Internal State Scale (ISS), Depression-Happiness Scale (D-HS), Manic Depressiveness Scale, Affective Self Rating Scale (ASRS), Hypomania Attitudes and the Positive Predictions Inventory (HAPPI) and Multidimensional Assessment of Thymic States (MAThyS). Moreover, Table I shows instruments that assess temperament, character and manic personality, the Temperament evaluation of Memphis, Pisa, Paris and San Diego (TEMPS-A) and the Affective Temperament Questionnaire (ATQ) and two rating scales developed to assess multiple psychiatric disorders, the

Concise Associated Symptoms Tracking Scale (CAST-SR) and the My Mood Monitor (M-3) Checklist.

Mood Disorders Questionnaire

The Mood Disorders Questionnaire (MDQ), developed by Hischfeld et al. in 2002 [7 23], consists of 13 dichotomous items reflecting DSM-IV criteria for a manic episode, 1 item inquiring whether the symptoms co-occurred, and 1 item in which the subject has to evaluate on a four-point scale the extent to which manic symptom led to significant and occupational dysfunction. In a review of Zimmermann and Galione [24] across all studies carried out in adults, the sensitivity of the MDQ was 0.61, specificity was 0.87, the positive predictive value was 0.58 and the negative predictive value was 0.88. Compared to studies using the MDQ for psychiatric outpatients, studies carried out in the general population found that it had much lower sensitivity and positive predictive value, and higher specificity and negative predictive value. The MDQ's sensitivity was higher in detecting bipolar I than bipolar II disorder (0.66 vs. 0.39). Lowering the threshold to identify cases markedly improved the MDQ's sensitivity, with only a modest reduction in specificity. Studies of the best symptom cut-off to identify cases produced inconsistent findings.

Based on currently available evidence, Zimmermann and Galione's [24] conclusions were that routine clinical use of the MDQ cannot be recommended because of the absence of studies simultaneously examining both the potential benefits (e.g. improved detection) and costs (e.g. overdiagnosis) of screening.

In the Italian validation study by Hardoy and Colleagues [25], the MDQ showed a good accuracy for bipolar or schizoaffective bipolar type disorders compared with patients with other axis I diagnoses or no diagnosis: the cut-off of 6 provided the best balance of sensitivity (0.76) and specificity (0.86). The accuracy for bipolar II disorders was sufficient but not excellent. The MDQ instrument was also recently validated in Chinese by Zhaoyu Gan et al. [26]. In this study, carried out in patients with mood disorders, the MDQ showed a good accuracy for bipolar disorder: the optimal cut-off was 4, with a sensitivity of 0.72 and a specificity of 0.73. Another recent Chinese study [27] on a large sample of patients treated for major depression found that one-fifth met criteria for bipolar disorders using the MINI Neuropsychiatric Interview. In this study, the MDQ had a poor ability to discriminate patients with and without bipolar disorder (sensitivity 0.30). In conclusion, given that the selection of the optimal cut-off requires a trade-off between sensitivity and specificity and that high values on both are desirable for use in clinical practice, the MDQ appears to possess a relatively high specificity and more limited sensitivity. In other words, the MDQ has a better performance in ruling out than in confirming a diagnosis of bipolar disorder.

Hypomania Checklist-32 (HCL-32)

The HCL-32 is a self-report questionnaire developed by Angst et al. in 2005 [28] and validated in different countries and languages (German, English, Swedish, Italian, Spanish, Polish, Chinese, Korean and Portuguese) that includes a list of (hypo)manic symptoms to be rated as present or absent with reference to a period of high state. It was originally designed to identify hypomanic symptoms in patients with major depressive disorder (MDD) in order to differentiate bipolar spectrum disorder (BSD) from MDD. The total HCL-32 score is obtained by adding up the positive responses to the 32 symptoms of hypomania. In the original validation study, comparing results from Italian and Swedish samples, the HCL distinguished BSD from MDD with a sensitivity of 0.80 and specificity of 0.51 [28]. In Italy, a further validation study performed by Carta and Colleagues [29] indicated good accuracy of HCL-32 as a screening instrument for BPD in a psychiatric setting, with a low rate of false negatives, and a fairly good degree of identification of bipolar II disorder. The recommended cut-off for discriminating bipolar disorders from other diagnoses (or no diagnosis) and for discriminating bipolar II disorder was 12.

Overall, evidence about this instrument indicates that its specificity is low because values around 0.50 indicate a performance no better that chance.

Bipolar Spectrum Diagnostic Scale (BSDS)

Ronald W. Pies [30] developed the BSDS to detect milder variants of bipolar disorder. The scale consists of two parts. The first part is a paragraph containing 19 statements describing many of the symptoms of bipolar disorder. For each statement, respondents are asked to place a checkmark if they believe that the statement applies to them. The second part of the BSDS is a single multiple-choice question asking respondents how well the paragraph describes them. The sensitivity for the bipolar I group was 0.75, while the sensitivity for the bipolar II/NOS group was 0.79. In comparison, only 15% of the unipolar subjects received false-positive screens from the BSDS, indicating a specificity of 0.85 [30]. The sensitivity of the BSDS proved to be similar for bipolar I disorder, bipolar II disorder and bipolar disorder NOS / cyclothymia in a large sample of 1,100 psychiatric outpatients [31]. A receiver operating curve (ROC) analysis indicated that cut-offs of 11 and 12 maximized the sum of sensitivity and specificity for the entire group of patients with bipolar disorder (AUC = 0.80, p < 0.001). The cut-off point associated with 90% sensitivity for the entire sam-

TABLE I.
Psychometric properties of manic/hypomanic symptom rating scales. The seven instruments described in the text are in boldface. *Proprietà psicometriche delle scale di valutazione dei sintomi maniacali/ipomaniacali. I sette strumenti descritti nel testo sono in grassetto.*

			ADMINISTRATION: SELF-REPORT				
Year	Instrument	Internal consistency	Concurrent/ discriminant validity	Factor analysis	Description/Aim	Inter-rater reliability	References
2011	Concise Associated Symptoms Tracking Scale Self-report (CAST-SR)	Alpha = 0.81 (17 item) Alpha = 0.78 (16 item)	The 5 CAST-SR domains correlated well with other standard measures of depressive severity and assessment of potential pre-cursors symptoms (BAI, HDRS, PDSQ, QIDS-C)	5 factors: Irritability Anxiety Mania Insomnia Panic	The CAST includes questions about irritability, anxiety, ma-nia, insomnia and panic do-mains thought to be associated with increased risk for suicide-related events and behaviours. The items in the CAST were designed to be rated using a Likert scale		54
2010	My Mood Monitor (M-3) Checklist	The M-3 bipolar module had a somewhat higher sen-sitivity (0.88; 95% CI, 0.77-0.95) but a lower specificity (0.70; 95% CI, 0.66-0.74). The anxiety module had a sensitivity of 0.82 (95% CI, 0.75-0.87) and a specificity of 0.78 (95% CI, 0.74-0.81), whereas the PTSD module had a sensitivity of 0.88 (95% CI, 0.74-0.96) and a specificity of 0.76 (95% CI, 0.73-0.80)			M-3 Checklist is a 23-item self-report symptom checklist that inquires whether during the past 2 weeks the patient experienced symptoms of major depressive disorder, generalized anxiety disorder, panic disorder, social anxiety disorder, PTSD and obsessive compulsive disorder. The M-3 also inquires about a lifetime history of symptoms of bipolar spectrum disorder. At the end of the symptom check-list, the M-3 poses 4 functional impairment questions. The M-3 is developed to screen for multiple psychiatric disor-ders in primary care		55
2009	Affective Temperament Questionnaire (ATQ)	Hyperthymia: alpha = 0.68 Cyclothymia: alpha = 0.83 Dysthymia: alpha = 0.81		3 factors: Hyperthymia Cyclothymia Dysthymia	Was designed to capture the es-sence of the criteria of Akiskal and Mallya (criteria defining 4 affective temperaments: hyper-thymic, irritable, cyclothymic, and dysthymic) in a self-rating form	Hyperthymia and cyclothymia were more prevalent among individuals with BP than among individuals with MDD or no history of a mood disorder. Dysthymia occurred at a relatively simi-lar rate among indi-viduals with MDD or BP	56

2008	Affective Self Rating Scale	The subscales for mania and depression showed high internal consistency with Cronbach's alphas of 0.89 for the depression subscale and 0.91 for the mania subscale	Depression subscore MADRS (r = 0.74) HIGH-C (r = 0.15) CGI-BP-D (r = 0.68) CGI-BP-M (r = -0.01) Mania subscore MADRS (r = 0.25) HIGH-C (r = 0.80) CGI-BP-D (r = 0.10) CGI-BP-M (r = 0.73)	4 factors	Measurement of intensity of current affective symptoms (depressive manic and mixed states)	57
2008	Multidimensional Assessment of Thymic States (MAThyS)	Alpha = 0.95	The MAThyS total score is moderately correlated of both the MADRS scale (depressive score; r = -0.45) and the MAS scale (manic score; r = 0.56)	5 factors: Emotional reactivity Motivation and psychomotor function Sensory perception Interpersonal communication Cognition	Discriminate between different sub-populations among patients suffering from bipolar disorders. The instrument is designed as a multi-dimensional assisted self-administered questionnaire comprising 20 items relating to individual states as perceived by patients for the preceding week	58
2006	Hypomania Attitudes and Positive Predictions Inventory (HAPPI; 61-item version)	Cronbach's alpha ranged from 0.83 for Increasing Activation to Avoid Failure and Grandiose Appraisals of Ideation to .90 for Social Self-Criticism. Internal consistency was Cronbach's .97 for the overall scale	HAPPI was significantly and positively related to prospective ISS Activation, Conflict and Depression. There was also a negative relationship between HAPPI and ISS Well-being	Factors: Social Self-Criticism Increasing Activation to Avoid Failure Success Activation & Triumph Over Fear Loss of Control Grandiose Appraisals of Ideation Regaining Autonomy	The HAPPI was developed to assess the multiple, extreme, and personalized beliefs key to an integrative cognitive model of bipolar depression and mood swings. The model postulated that bipolar symptoms are developed and maintained by interpreting physiological, affective, and cognitive changes to internal states, and perceived behavioral changes, as having extreme personal meaning. These appraisals are multiple, positive and negative, and can therefore be conflicting	59, 60

(continues)

(Table I follows)

			ADMINISTRATION: SELF-REPORT				
Year	Instrument	Internal consistency	Concurrent/ discriminant validity	Factor analysis	Description/Aim	Inter-rater reliability	References
2005	Hypomania Checklist-32 (HCL-32)	Alpha = 0.82 (Italian sample) Alpha = 0.86 (Swedish sample)		2 factors: Active/elated Risk-taking/irritable	The primary goal of the HCL-32 is to identify hypomanic components in patients with MDD in order to help the clinician to diagnose BP-II and other BP spectrum disorders presenting in psychiatric and general medical practice		28
2005	Bipolar Spectrum Diagnostic Scale (BSDS)	Sensitivity: 0.75 for the bipolar I group 0.79 for the bipolar II/NOS. 15% of the unipolar subjects received false-positive screens from the BSDS, indicating a specificity of 0.85			Designed to detect the milder portions of the bipolar spectrum in outpatients. The final version is composed of two parts. The first part is a paragraph containing 19 positively valenced sentences describing many of the symptoms of bipolar disorder. The second part is one simple multiple-choice question, asking subjects to rate how well the story describes them overall		30
2005	Temperament evaluation of Memphis, Pisa, Paris and San Diego (TEMPS-A)	Alpha for the four scales: Cyclothymic 0.88 Irritable 0.84 Hyperthymic 0.81 Dysthymic 0.76		4 factors: Cyclothymic Irritable Hyperthymic Dysthymic			61
2002	Mood Spectrum Self-Report (MOODS-SR) Available in the lifetime and last-month format			11 factors: Psychomotor activation Mixed instability Spirituality/mysticism/psychoticism Mixed irritability Euphoria Depressive mood Psychomotor retardation Suicidality Drug/illness related depression Psychotic features Neurovegetative symptoms	The SCI-MOODS and the MOODS-SR includes 161 items coded as present/absent, for one or more periods of at least 3 to 5 days. Items explore depressive and manic mood, energy, and cognition, and disturbances in rhythmicity		37

Year	Instrument	Reliability	Validity	Factors	Description	Ref
2000	Mood Disorder Questionnaire (MDQ)	Alpha = 0.90			The Mood Disorder Questionnaire is a self-report, single-page inventory that screens for a lifetime history of a manic or hypomanic syndrome by including 13 yes/no items derived from both the DSM-IV criteria and clinical experience	7
1998	Manic Depressiveness Scale	The Cronbach a for the depressive items was 0.63 and for the mania items was 0.56	The subscales distinguish well between the patients and the control group		Detection of people who at some point in their lives had experienced behaviours or actions typical of bipolar disorders. This type of scale allows not only the assessment of past or present experiences with this disorder but can also be used as an indicator of cyclothymic syndromes or attenuated forms	62
1998	The Depression-Happiness Scale	Alpha = 0.88	Higher scores on the D–H S were associated with higher scores on the OHI, $r = .59$, $p < .001$, and lower scores on the BDI, $r = -.75$, $p < .001$, confirming the construct validity of the scale		This is a self-report scale which contains 25 items representing a mix of affective, cognitive, and bodily experiences. Higher scores on the scale indicate a higher frequency of positive thoughts, feelings, and bodily experiences and a lower frequency of negative thoughts, feelings, and bodily experiences	63
1997	Altman Self-rating Mania Scale	Alpha: mania = 0.79 psychosis = 0.65 irritability = 0.65	The Pearson correlation coefficient between ASRM mania subscale scores and MRS total scores was $r = .718$ ($p < .001$) and between ASRM mania subscale scores and CARS-M mania subscale scores $r = .766$ ($p < .001$)	3 factors: Mania Psychosis Irritability	Brief self-rating mania scale, compatible with DSM-IV criteria, used to measure the presence and severity of manic symptoms for research or clinical purposes	19

(continues)

(Table I follows)

ADMINISTRATION: SELF-REPORT

Year	Instrument	Internal consistency	Concurrent/ discriminant validity	Factor analysis	Description/Aim	Inter-rater reliability	References
1992	Self-Report Manic Inventory (SRMI)	Alpha = 0.94	The discrimination analysis shows that the questionnaire seems to differentiate well between manic and non-manic subjects (71% of the samples are well classified) Test-retest reliability: (r = 0.93 for manic subjects and 0.73 for the whole sample	2 factors: Energized Dysphoria Hedonistic Euphoria	Includes the symptoms of mania as described in the DSM-III-R and also in some authoritative works together with the authors' experience. The 47 items contained in the scale are answered through items with true/false options. An insight question is added to these 47 items		46
1991	Internal State Scale	Alpha: from 0.81 to 0.92 on the four factors	The activation subscale gave a strong correlation with the Manic Rating Scale by Young (r = 0.60). The Depression Index correlated with that of Hamilton (r = 0.84) as did that of Well-being (r = 0.73)	4 factors: Depression Index Wellbeing Activation Perception of Conflict	Instrument for assessment of manic and depressive symptoms by patients and their families		64
1973	Visual Analogue Mood Scale		VAMS correlated with the SDS and several subscales of the Clyde Mood Scale The VAMS digit-symbol combination was able to distinguish patients with affective disorder from others, better than other tests used		The VAMS is a rectangular card 100 mm by 35 mm on which the following instruction is printed: 'How is your mood right now? A mark on the line toward the left represents your worst mood, toward the right, your best.' The VAMS score is determined by measuring the distance in millimetres from the left end of the card to the patient's mark		65

		ADMINISTRATION: INTERVIEW-BASED			
2011	Concise Associated Symptoms Tracking Scale Clinical rating (CAST-C)	CAST-C: Alpha = 0.80 (17 item) Alpha = 0.77 (16 item)	The 5 CAST-C domains correlated well with other standard measures of depressive severity and assessment of potential precursors symptoms (BAI, HDRS, PDSQ, QIDS-C)	5 factors: Irritability Anxiety Mania Insomnia Panic	The CAST includes questions about irritability, anxiety, mania, insomnia and panic domains thought to be associated with increased risk for suicide-related events and behaviors. The items in the CAST were designed to be rated using a Likert scale	54
2009	Biological Rhythms Interview of Assessment in Neuropsychiatry (BRIAN)	Cronbach's alpha = 0.93	Total BRIAN scores were highly correlated with the global Pittsburgh Sleep Quality Index score (rho = 0.77, p < 0.001) Highly significant differences between the clinical and control group were found for the whole scale (U = 925.5, p < 0.001)	3 factor solution: Sleep/social Rhythm factor Activity factor Feeding factor	Designed to assess biological rhythms in the clinical setting	66
2007	Bipolar Inventory of Symptoms Scale (BISS)	Cronbach's alpha for the total scale was 0.90, 0.93 for the depression subscale and 0.92 for the mania subscale. Internal consistency ranged from 0.69 to 0.93 for all symptom clusters except somatic complaints (0.39)	Total BISS and manic subscale scores were significantly associated with GAF scores (-0.52, -0.48). Irritability cluster scores were significantly associated with CGI-BD and GAF scores (0.63, -0.56)as were psychosis cluster scores (0.56, - 0.67). No depression cluster or the depression subscale was significantly associated with either CGI-BD or GAF Discriminant validity: the BISS total score was significantly higher in each syndromal mood state (depressed, mixed, and manic/hypomanic) than in recovered, but did not differ between syndromes	5 factors: Mania Depression Irritability Anxiety psychosis	The BISS is designed to comprehensively assess and quantitate the symptomatology of BDs, including bipolar II and spectrum patients, and to provide a sensitive instrument for the detection of treatment effects	67

(Table I follows)

			ADMINISTRATION: INTERVIEW-BASED				
Year	Instrument	Internal consistency	Concurrent/ discriminant validity	Factor analysis	Description/Aim	Inter-rater reliability	References
1999	Structured Clinical Interview for Mood Spectrum (SCI-MOODS)	Internal consistency for the seven domains ranged between 0.72 and 0.92			Designed to evaluate the life-time presence/absence of the DSM-IV core symptoms of depression and mania, atypical symptoms, subthreshold manifestations and behavioural traits that arise as a mean of coping with mood symptoms		36
1999	Coping inventory for prodromes of mania (CIPM)	Alpha: Stimulation reduction = 0.77 Problem-directed coping = 0.85 Seeking professional help = 0.53 Denial or blame = 0.70		4 factors: Stimulation reduction Problem-directed coping Seeking professional help Denial or blame	Instrument aimed at assessing how manic depressive sufferers dealt with their prodromes of mania		69
1998	Temperament evaluation of Memphis, Pisa, Paris and San Diego (TEMPS)	Alpha for the four scales: Depressive 0.85 Hyperthymic 0.86 Cyclothymic 0.94 Irritable 0.88		Four factors: depressive hyperthymic cyclothymic irritable			70
1994	Clinician-Administered Rating Scale for Mania (CARS-M)	Alpha = 0.93	Manic Rating Scale (r = 0.94). The sensitivity of the scale is very high, differentiating between diagnostic groups (p < 0.001)	2 factors. Mania (10 items). Psychoticism (5 items)	Scale for the assessment and quantification of Mania; contains 15 items rated on a Likert scale from 1 to 5, except for one which goes from 1 to 4 It is a symptomatic scale also including information culled by other members of the hospital unit or family members		71
1988	Manchester Nurse Rating Scale for Mania (MNRS-M)	All the individual items of the MNRS-M were significantly correlated with the total mania score	Correlation with Shopsin Global Mania Rating was 0.65 Correlation with YMRS was 0.79		The scale, administered from the nursing staff, was designed for the daily rating of manic ward behaviours	The product-moment correlation coefficients (r) obtained ranged from 0.813 to 0.993 (p < 0.05)	72
1986	Hypomanic Personality Scale		Test-retest reliability (15 weeks): 0.81		A 48-item true–false scale measuring hyperactive, ambitious, and exhibitionistic behaviours as well as feelings of euphoria and flights of thoughts		73

Year	Instrument	Reliability	Validity	Factor	Description	Inter-rater reliability	Ref
1978	Bech-Rafaelsen's mania scale				Clinician interview assessing the current manic symptoms; comprises 11 items defined against a five-point scale	inter-rater reliability is high (r = 0.80-0.95 for four raters)	43
1978	Young Mania Rating Scale (YMRS) Clinical interview		The total scores on the YMRS correlated highly with the global rating (0.88) and the Petterson Scale (0.89). The correlation with the Beige scale, although of a lower magnitude (0.71), was acceptable		The Young Mania Rating Scale consists of eleven items, each with five explicitly defined grades of severity. The choice of items was made on the basis of published descriptions of the core symptoms of the manic phase of bipolar affective disorder and includes those abnormalities which were felt to exist over the entire range of illness from mild to severe	The correlation between the ratings of two physicians was 0.93 for the total YMRS score and ranged from 0.66 for item 9, disruptive-aggressive behaviour, to 0.95 for item 4, sleep. All correlations were significant at the 0.001 level	39
ADMINISTRATION: OTHER							
2010	The Interactive Computer Interview for Mania (ICI-M) Computer-administered interview	Alpha = 0.82			Computer-administered interview that both presents probes designed to elicit information about the presence and severity of symptoms and utilizes a scoring algorithm to select follow-up questions and rate subject responses in accordance with rating scale anchor points. The goal of the ICI is to provide a standard comparator that can be used to enhance the sensitivity and consistency of human raters by providing on-going feedback on the concordance of their ratings with the ICI ratings and identify those in need of specific remediation during the course of study operations	The intraclass correlation coefficients was 0.91 between the ICI-M and an expert consensus rating	74
2010	Observer-Rated Scale for Mania (ORSM)	Alpha = 0.89 Test-retest reliability. r = 0.76-0.89	Correlations with DSM-IV were high with a Pearson's correlation coefficient ranging from of .74 and .70 (p < .01). Similarly, Pearson's correlation coefficient for the YMRS ranges from .75 and .76 (p < .01) and for the MSS ranges from .65 and .62 (p < .05)	The three-factor solution accounted for 70.5% of the variance. Euphoric mania Instable mania Psychotic mania	A mania rating scale that can be used by individuals who are in close contact with the patient in order to assess mania and to determine its severity		75

ple of patients with bipolar disorder was 8. At this cut-off, the specificity of the scale was 51.1% with a positive predictive value of 16.0%. The authors compared patients with and without bipolar disorder on each of the BSDS symptom items. The odds ratios were higher for the items assessing hypomanic / manic symptoms than those assessing depressive symptoms, although the performance of a subscale composed only of the hypomania / mania items was nearly identical to that of the entire scale (AUC = 0.81, p < 0.001).

Zimmermann et al. [31] concluded that this instrument is excellent at ruling out a diagnosis of bipolar disorder; however, the low positive predictive value indicates that it is not good at confirming diagnosis. These data raise questions about the use of the BSDS as a screening measure in routine clinical psychiatric practice.

Cross-validations of MDQ, HCL-32 and BSDS

In China, the optimal cut-off for the MDQ was determined as 4 symptoms endorsed and that for the HCL-32 11 symptoms. The administration of the HCL-32 coupled with the collection of family history for bipolar disorder proved to be more efficient than the MDQ to detect bipolar disorder [32].

In Italy, Carta et al. [29] examined the accuracy of HCL-32 in detecting bipolar I and II disorders compared with the MDQ. When the balance between sensitivity and specificity was considered, the MDQ had better performance in detecting bipolar disorders, although the HCL-32 was better in discriminating bipolar-II disorder.

In Spain, Vieta et al. [33] compared the discriminative capacity of the two instruments, using the HCL-32 sensitivity and specificity indices and verifying that the confidence intervals of the Spanish version of the MDQ contained the value of the HCL-32 indices. The sensitivity of the HCL-32 was 0.85 with a specificity of 0.79. The confidence intervals for the sensitivity and specificity of the MDQ were 95% CI (0.51, 0.69) and 95% CI (0.94, 0.99), respectively. Because the sensitivity and specificity values of the HCL-32 fell outside the confidence intervals of the MDQ, the authors concluded that the HCL-32 had higher sensitivity but less specificity than the MDQ. This procedure, however, appears questionable because the authors should have tested the statistical difference between the areas under the curve of the two instruments to draw a conclusion about their comparative performance. In Poland, Rybakowski et al. [34] reported that hypomanic symptoms exceeding cut-off criteria for bipolarity by HCL-32 were present in 37.5% of patients and, by MDQ, in 20% of patients. The percentage of patients with treatment-resistant depression was significantly higher both in patients screening positive on HCL-32 and in those screening positive on MDQ compared to those screening

negative, providing additional evidence of the discriminant capacity of the two instruments.

In an English study, the HCL-32 performed better than the Bipolar Spectrum Diagnostic Scale as a means of identifying bipolar disorder in primary care, although the positive predictive values of both instruments were relatively low [35].

In summary, these studies suggest that the best screening instrument is the HCL-32, because of its higher sensitivity and ability to discriminate bipolar I from bipolar II disorder.

The Mood Spectrum (MOODS-SR)

The MOODS-SR is a self-report instrument consisting of 161 dichotomous items designed to provide a careful assessment of depressive and manic/hypomanic features that may have occurred for at least 3-5 days during the individual's lifespan. The instrument was first developed in English and Italian in parallel by a panel of experts as a structured interview [36] and then as a self-report questionnaire [37]. In both formats it has been shown to have excellent psychometric properties. A recently published paper reported the results of a classification tree analysis of a pooled dataset of 1158 patients with bipolar disorders (I, II and NOS) or unipolar depression participating in 5 studies [38] who were administered the MOODS-SR.

Using 11 dimensions derived from factor analyses that characterize the manic and the depressive side of the mood spectrum (psychomotor activation, mixed instability, spirituality/mysticism/psychoticism, mixed irritability, euphoria, depressive mood, psychomotor retardation, suicidality, drug/illness-related depression, psychotic spectrum features, neurovegetative symptoms), the authors determined that only 4 of these dimensions (psychomotor activation, mixed instability, suicidality and euphoria, comprising overall 33 dichotomous items, see Table II) are needed to stratify the sample into subgroups with a differential risk of bipolarity, and identified cut-off scores for each dimension to be used in clinical practice. This paper deserves mention because it provides an empirical confirmation of the role of psychomotor activation as the key feature of bipolar disorder. This finding is in line with the recent changes proposed by the DSM-V, which revised criterion A to include increased energy/activity as a core symptom of mania.

Interview-based instruments

Among observer-rated instruments, the nurses' rating scales were the first instruments to appear, followed by specific instruments for application by clinicians.

Table I lists 12 interview-based instruments, an observer-rated instrument and a computer-administered interview.

TABLE II.
MOODS-SR items discriminating unipolar from bipolar patients. *Item del MOODS-SR che discriminano tra pazienti bipolari e unipolari.*

PSYCHOMOTOR ACTIVATION
68. Urge to communicate
69. Desire to reconnect with people
70. Talkative
71. Noisy
72. Racing thoughts
73. Too many thoughts at once
74. Shifting interests
76. Assertive
77. Vigorous
78. Very impatient
79. Constantly active
125. Irresponsible
136. More energetic with less sleep
143. Not tired even without sleeping

MIXED INSTABILITY
46. Frequently changing: job, residence, friends, hobbies
47. Risk taking
55. Irritable or elevated mood when you were abusing alcohol
56. Irritable or elevated mood when you increased your use alcohol
126. Made important decisions very rapidly
128. Tended to ignore everyday rules and social etiquette
157. More interested in sex
158. Changed sexual partners

EUPHORIA
29. Persistently good or high
30. High sense of humor and irony
31. Even the smallest thing could you very enthusiastic
32. Liked to make puns or plays on words
33. Making a lot of jokes

SUICIDALITY
102. Life is not worth living
103. Wishing not to wake up in the morning
104. Want to die or hurt yourself
105. Specific plan to hurt or kill yourself
106. Suicide attempt
107. Suicide attempt requiring medical attention

Five of these instruments are designed to assess the manic pole of the bipolar disorder (Clinician-Administered Rating Scale for Mania (CARS-M), Observer-Rated Scale for Mania (ORSM), Interactive Computer Interview for Mania (ICI-M), Young Mania Rating Scale (YMRS) and Bech-Rafaelsen Mania Scale (BRMES)) and four to assess both poles (Bipolar Inventory of Symptoms Scale (BISS), Brief Bipolar Disorder Symptom Scale (BDSS), Manchester Nurse Rating Scale for Mania (MNRS-M) and the Structured Clinical Interview for Mood Spectrum (SCI-MOODS). Moreover, in Table I two instruments assessing temperament, character and manic personality (the Temperament evaluation of Memphis, Pisa, Paris and San Diego [TEMPS] and the Hypomanic Personality Scale), a rating scale assessing multiple psychiatric disorders (Concise Associated Symptoms Tracking Scale; CAST-C), an instrument designed to assess biological rhythms in the clinical setting (Biological Rhythms Interview of Assessment in Neuropsychiatry; BRIAN) and an instrument that examines the coping strategies used by manic depressive patients during the prodromal phase of their manic episodes (Coping Inventory for Prodromes of Mania [CIPM]) are presented.

Below we list the characteristics of the two most widely used scales in clinical trials.

The Young Mania Rating Scale (YMRS)

This scale, developed by Young in 1978 [39], includes 11 items and is used to assess disease severity in patients already diagnosed with mania. It is intended to be administered by a trained clinician who assigns a severity rating on a Likert scale for each item based on a personal interview. The total score ranges from 0 to 56. The scale is based on the patient's subjective report of his/her clinical condition over the previous 48 hours that typically takes 15-30 minutes to administer. Items can be rated by querying the patients or from direct observation, and encompass elevated mood, increased motor activity, sexual interest, sleep, irritability, speech, language/thought disorder, content, disruptive/aggressive behaviour, appearance and insight. It is the most used outcome measurement in clinical trials and longitudinal naturalistic studies. A cut-off on the total YMRS score < 4 was suggested by Berk [40] to denote complete remission. Gonzalez-Pinto [41] used a cut-off > 20 for acute mania and Benvenuti et al. [42] used a cut-off of ≥ 10 to define a manic/hypomanic switch in patients with unipolar depression.

The Bech-Rafaelsen mania scale (BRMAS)

The BRMAS [43 44] is used to assess current manic symptoms and takes 15-30 minutes to administer. The 11 items are rated on a 5-point scale and each rating has very specific anchor points that facilitate the rating. The items explore

motor activity, verbal activity, flight of thoughts, voice/noise level, hostility/destructiveness, feelings of well being, self-esteem, contact with others, sleep changes, sexual interests, and work activities, similarly to YMRS, but do not assess insight and appearance. This scale has been frequently used as an outcome measure in clinical trials for more than 30 years. Studies of the internal validity of the BRMAS have demonstrated that the simple sum of the 11 items of the scale is a sufficient statistic for the assessment of the severity of manic states. Both factor analysis and latent structure analysis (the Rasch analysis) have been used to demonstrate that the scale is unidimensional. The total score of the BRMAS has been standardized so that scores between 15 and 20 indicate mild hypomania, scores between 20 and 27 indicate moderate mania, and scores $\geq$ 28 indicate severe mania. The interrater reliability has been found to be high in a number of studies conducted in various countries [45].

Discussion

Traditionally, observer-rated scales have been used to measure manic states and self-rating scales have been developed only more recently. The latter have the advantage of being able to assess the patient's internal states and avoiding possible misinterpretation of clinicians, although some authors argued that their subjective nature makes them at risk of exaggeration or understatement of symptoms and non-standard interpretations of the meaning of the questions [46].

In the experimental research context, self-report and interview-based instruments are commonly utilized for preselected patients with mood disorders who have well-established diagnoses to assess treatment outcomes in terms of response/remission [30 47-49]. In contrast, in routine clinical practice no one would argue that rating scales eliminate the need for a competent psychiatric evaluation, considering that there are no 'special questions' on the most widely used scales that are unfamiliar to a competent clinician [50]. Nonetheless, rating scales may be very useful in clinical practice when it comes to making sure that specific and standardized questions (e.g. suicidal ideation) are consistently asked and recorded. Moreover, evidence that the administration of rating scales might improve the efficiency of diagnostic evaluation outside clinical trials (characterized by well-defined inclusion/exclusion criteria) is still controversial [30 28 51]. In particular, two potential negative consequences have been commonly reported with the systematic assessment provided by self-rating scales for mania in clinical settings. On the one hand, the sensitivity of several instruments is around 60-65%, and clinicians who rely on screening scales that use the first stage in a two-stage process for di-

agnosing bipolar disorders are at high risk of missing the correct diagnosis in approximately one-third of patients. On the other hand, the positive predictive value of such instruments is often inadequate, raising the possibility of an over-diagnosis of bipolar disorder, if no more valid and comprehensive diagnostic assessment tools are subsequently provided.

In developing a broad-based screening measure for multiple psychiatric disorders, Zimmermann and Mattia [52] recommended that a cut-off resulting in diagnostic sensitivity of 90% and a correspondingly high negative predictive value be chosen when using an instrument in clinical practice. With high negative predictive value, the clinician can reliably assume that when the test indicates that the disorder is not present, inquiring about the disorder's symptoms is pointless. Our review of self-report instruments indicates that this sensitivity level is not achieved by the MDQ, the HCL-32 or the BDRS and, if it is achieved, this happens at the cost of very high false positive rates. In general, the trade-off between sensitivity and specificity depends on the disease and the specific purpose of the screening. In the case of bipolar disorder, early identification of patients suffering from this condition is as important as excluding this diagnosis to develop a suitable treatment strategy.

The psychometric properties of the instruments reviewed suggest that the MDQ might be useful for screening patients presenting with recurrent depression or anxiety to rule out the presence of bipolar disorder in psychiatric clinical settings and primary care. The HCL-32, which is more sensitive than the MDQ, might be used to screen potential cases to be further investigated with a diagnostic interview. On this note, it should be emphasized that other elements such as family history, age of onset of symptoms, course of symptoms and previous response to medication play a key role in the diagnostic process.

Considering the MOODS-SR, the instrument is relatively long, which makes it more suitable for research purposes than for routine clinical use. Still, the 33 items exploring the key features discriminating bipolar disorder from unipolar depression seem to be promising as a stand-alone screening instrument to detect the presence of manic/hypomanic features lasting at least 3-5 days in the lifetime. However, to date no study has provided evidence of the psychometric properties of this subset of items.

The 5-item Altman Mania Rating Scale appears to be useful for monitoring the longitudinal course of mania/hypomania symptoms for research and clinical purposes and generates results similar to those of other longitudinal studies of bipolar disorder that use traditional retrospective, clinician-gathered mood data [21].

Regarding the interview-based instruments, the YMRS and the BRMAS have a similar coverage, although the

YMRS is to be preferred over the BRMAS because it includes an item on insight. In fact, when the severity of symptoms affects the level of insight, the reliability of the assessment may be compromised.

In conclusion, although no instrument can replace the need for accurate clinical diagnosis based on patient history, we argue that the increasing use of self-report instruments to screen bipolar disorder in high-risk patients presenting with depression or to monitor mania/hypomania symptoms over time may contribute to increasing the use of routine standardized assessment. Measurement-based treatment as the standard of care has the potential to transform psychiatric practice, move psychiatry into the mainstream of medicine, facilitate integration between primary care and mental health services [53] and ultimately improve the quality of care for patients with psychiatric illness [54].

References

1. Merikangas KR, Jin R, He JP, et al. *Prevalence and Correlates of Bipolar Spectrum Disorder in the World Mental Health Survey Initiative*. Arch Gen Psychiatry 2011;68:241.

2. Hantouche EG, Akiskal HS, Lancrenon S, et al. *Systematic clinical methodology for validating bipolar-II disorder: data in mid-stream from a French national multi-site study (EPIDEP)*. J Affect Disord 1998;50:163-73.

3. Ghaemi SN, Boiman EE, Goodwin FK. *Diagnosing bipolar disorder and the effect of antidepressants: a naturalistic study*. J Clin Psychiatry 2000;61:804-8.

4. Ghaemi SN, Ko JY, Goodwin FK. *''Cade's disease'' and beyond: misdiagnosis, antidepressant use, and a proposed definition for bipolar spectrum disorder*. Can J Psychiatry 2002;47:125-34.

5. Zimmerman M, Ruggero CJ, Chelminski I, et al. *Is bipolar disorder overdiagnosed?* J Clin Psychiatry 2008;69:935-40.

6. Lish J, Dime-Meenan S, Whybrow P, et al. *The National Depressive and Manic-depression Association (DMDA) survey of bipolar members*. J Affect Disord 1994;31:281-94.

7. Hirschfeld RM, Williams JB, Spitzer RL, et al. *Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire*. Am J Psychiatry 2000;157:1873-5.

8. Birnbaum HG, Shi L, Dial E, et al. *Economic consequences of not recognizing bipolar disorder patients: a cross-sectional descriptive analysis*. J Clin Psychiatry 2003;64:1201-9.

9. Shi L, Thiebaud P, McCombs JS. *The impact of unrecognized bipolar disorders for patients treated for depression with antidepressants in the fee-for-services California Medicaid (Medi-Cal) program*. J Affect Disord 2004;82:373-83.

10. Matza LS, Rajagopalan KS, Thompson CL, et al. *Misdiagnosed patients with bipolar disorder: comorbidities, treatment patterns, and direct treatment costs*. J Clin Psychiatry 2005;66:1432-40.

11. Judd L, Akiskal H, Schettler P, et al. *The long-term natural history of the weekly symptomatic status of bipolar I disorder*. Arch Gen Psychiatry 2002;59:530-7.

12. Cassano GB, Dell'Osso L, Frank E, et al. *The bipolar spectrum: a clinical reality in search of diagnostic criteria and an assessment methodology*. J Affect Disord 1999;54:319-28.

13. Bowden CL. *Strategies to reduce misdiagnosis of bipolar depression*. Psychiatr Serv 2001;52:51-5.

14. Isometsa E, Suominen K, Mantere O, et al. *The Mood Disorder Questionnaire improves recognition of bipolar disorder in psychiatric care*. BMC Psychiatry 2003;10:3-8.

15. Hirschfeld RM, Vornik LA. *Recognition and diagnosis of bipolar disorder*. J Clin Psychiatry 2004;65(Suppl 15):5-9.

16. Yatham LN. *Diagnosis and management of patients with bipolar II disorder*. J Clin Psychiatry 2005;66(Suppl 1):13-7.

17. Livianos-Aldana L, Rojo-Moreno L. *Rating and quantification of manic syndromes*. Acta Psychiatr Scand Suppl 2001;409:2-33.

18. Picardi A. *Rating scales in bipolar disorder*. Curr Opin Psychiatry 2009;22:42-9.

19. Altman EG, Hedeker D, Peterson JL, et al. *The Altman Self-Rating Mania Scale*. Biol Psychiatry 1997;42:948-55.

20. Altman E. *Rating scales for mania: is self-rating reliable?* J Affect Disord 1998;50:283-6.

21. Bopp JM, Miklowitz DJ, Goodwin GM, et al. *The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study*. Bipolar Disord 2010;12:327-34.

22. Miklowitz DJ, Price J, Holmes EA, et al. *Facilitated Integrated Mood Management for adults with bipolar disorder*. Bipolar Disord 2012;14:185-97.

23. Hirschfeld RM. *The Mood Disorder Questionnaire: A Simple, Patient-Rated Screening Instrument for Bipolar Disorder*. Prim Care Companion J Clin Psychiatry 2002;4:9-11.

24. Zimmerman M, Galione JN. *Screening for bipolar disorder with the Mood Disorders Questionnaire: a review*. Harv Rev Psychiatry 2011;19:219-28.

25. Hardoy MC, Cadeddu M, Murru A, et al. *Validation of the Italian version of the ''Mood Disorder Questionnaire'' for the screening of bipolar disorders*. Clin Pract Epidemol Ment Health 2005;21:1-8.

26. Gan Z, Han Z, Li K, et al. *Validation of the Chinese version of the "Mood Disorder Questionnaire" for screening bipolar disorder among patients with a current depressive episode*. BMC Psychiatry 2012;31:12-8.

27. Hu C, Xiang YT, Wang G, et al. *Screening for bipolar disorder with the Mood Disorders Questionnaire in patients diagnosed as major depressive disorder – The experience in China*. J Affect Disord 2012;141:40-6.

28. Angst J, Adolfsson R, Benazzi F, et al. *The HCL-32: towards a self-assessment tool for hypomanic symptoms in outpatients*. J Affect Disord 2005;84:217-33.

29. Carta MG, Hardoy MC, Cadeddu M, et al. *The accuracy of the Italian version of the Hypomania Checklist (HCL-32) for*

the screening of bipolar disorders and comparison with the Mood Disorder Questionnaire (MDQ) in a clinical sample. Clin Pract Epidemiol Ment Health 2006;8:2.

30 Ghaemi SN, Miller CJ, Berv DA, et al. Sensitivity and specificity of a new bipolar spectrum diagnostic scale. J Affect Disord 2005;84:273-7.

31 Zimmerman M, Galione JN, Chelminski I, et al. Performance of the Bipolar Spectrum Diagnostic Scale in psychiatric outpatients. Bipolar Disord 2010;12:528-38.

32 Poon Y, Chung KF, Tso KC, et al. The use of Mood Disorder Questionnaire, Hypomania Checklist-32 and clinical predictors for screening previously unrecognised bipolar disorder in a general psychiatric setting. Psychiatry Res 2012;195:111-7.

33 Vieta E, Sánchez-Moreno J, Bulbena A, 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). J Affect Disord 2007;101:43-55.

34 Rybakowski JK, Dudek D, Pawlowski T, et al. Use of the Hypomania Checklist-32 and the Mood Disorder Questionnaire for detecting bipolarity in 1,051 patients with major depressive disorder. Eur Psychiatry 2012;27:577-81.

35 Smith DJ, Griffiths E, Kelly M, et al. Unrecognized bipolar disorder in primary care patients with depression. Br J Psychiatry 2011;199:49-56.

36 Fagiolini A, Dell'Osso L, Pini S, et al. Validity and reliability of the new instrument for assessing mood symptomatology: the Structured Clinical Interview for Mood Spectrum (SCI-MOODS). Int J Meth Psychiatr Res 1999;8:75-86.

37 Dell'Osso L, Armani A, Rucci P, et al. Measuring mood spectrum: comparison of interview (SCI-MOODS) and self-report (MOODS-SR) instruments. Compr. Psychiatry 2002;43:69-73.

38 Cassano GB, Rucci P, Benvenuti A, et al. The role of psychomotor activation in discriminating unipolar from bipolar disorders: a classification-tree analysis. J Clin Psychiatry 2012;73:22-8.

39 Young RC, Biggs JT, Ziegler VE, et al. A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry 1978;133:429-35.

40 Berk M, Ng F, Wang WV, et al. The empirical redefinition of the psychometric criteria for remission in bipolar disorder. J Affect Disord 2008;106:153-8.

41 González-Pinto A, Galán J, Martín-Carrasco M, et al. Anxiety as a marker of severity in acute mania. Acta Psychiatr Scand 2012;126:351-5.

42 Benvenuti A, Rucci P, Miniati M, et al. Treatment-emergent mania/hypomania in unipolar patients. Bipolar Disord 2008;10:726-32.

43 Bech P, Bolwig TG, Kramp P, et al. The Bech-Rafaelsen Mania Scale and the Hamilton Depression Scale. Acta Psychiatr Scand 1979;59:420-30.

44 Bech P, Kastrup M, Rafaelsen OJ. Mini-compendium of rating scales for states of anxiety depression mania schizophre-

nia with corresponding DSM-III syndromes. Acta Psychiatr Scand Suppl 1986;326:1-37.

45 Bech P. The Bech-Rafaelsen Melancholia Scale (MES) in clinical trials of therapies in depressive disorders: a 20-year review of its use as outcome measure. Acta Psychiatr Scand 2002;106:252-64.

46 Shugar G, Schertzer S, Di Gasbarro J. Development, use and factor analysis of a self-report inventory for mania. Compr Psychiatry 1992;33:325-31.

47 Weber Rouget B, Gervasoni N, Dubuis V, et al. Screening for bipolar disorders using a French version of the Mood Disorder Questionnaire (MDQ). J Affect Disord 2005;88:103-8.

48 Twiss J, Jones S, Anderson I. Validation of the Mood Disorder Questionnaire for screening for bipolar disorder in a UK sample. J Affect Disord 2008;110:180-4.

49 Zimmerman M. Misuse of the Mood Disorders Questionnaire as a case-finding measure and a critique of the concept of using a screening scale for bipolar disorder in psychiatric practice. Bipolar Disord 2012;14:127-34.

50 Parker G, Fletcher K, Barrett M, et al. Screening for bipolar disorder: the utility and comparative properties of the MSS and MDQ measures. J Affect Disord 2008;109:83-9.

51 Zimmerman M, Mattia JI. A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire (PDSQ). Arch Gen Psychiatry 2001;58:787-94.

52 Harding JK, Rush AJ, Arbuckle M, et al. Measurement-based care in psychiatric practice: a policy framework for implementation. J Clin Psychiatry 2011;72:1136-43.

53 Rucci P, Piazza A, Menchetti M, et al. Integration between Primary Care and Mental Health Services in Italy: determinants of referral and stepped care. Int J Family Med 2012, Article ID 507464. doi:10.1155/2012/507464.

54 Trivedi MH, Wisniewski SR, Morris DW, et al. Concise Associated Symptoms Tracking scale: a brief self-report and clinician rating of symptoms associated with suicidality. J Clin Psychiatry 2011;72:765-74.

55 Gaynes BN, DeVeaugh-Geiss J, Weir S, et al. Feasibility and Diagnostic Validity of the M-3 Checklist: A Brief, Self-Rated Screen for Depressive, Bipolar, Anxiety, and Post-Traumatic Stress Disorders in Primary Care. Ann Farm Med 2010;8:160-9.

56 Light KJ, Joyce PR, Frampton CM. Description and validation of the Affective Temperament Questionnaire. Compr Psychiatry 2009;50:477-84.

57 Adler M, Liberg B, Andersson S, et al. Development and validation of a self rating scale for manic, depressive and mixed affective states. Nord J Psychiatry 2008;62:130.

58 Henry C, M'Bailara K, Mathieu F, et al. Construction and validation of a dimensional scale exploring mood disorders: MAThyS (Multidimensional Assessment of Thymic States). BMC Psychiatry 2008;19:82.

59 Dodd AL, Mansell W, Sadhnani V, et al. Principal components analysis of the Hypomanic Attitudes and Positive Predictions Inventory and associations with measures of

personality, cognitive style and analogue symptoms in a student sample. Behav Cogn Psychother 2010;38:15-33.

60   Mansell W. *The Hypomanic Attitudes and Positive Predictions Inventory (HAPPI): A pilot study to select cognitions that are elevated in individuals with bipolar disorder compared to non-clinical controls.* Behav Cogn Psychother 2006;34:467-76.

61   Akiskal HS, Akiskal KK, Haykal RF, et al. *TEMPS-A: progress towards validation of a self-rated clinical version of the Temperament Evaluation of the Memphis, Pisa, Paris, and San Diego Autoquestionnaire.* J Affect Disord 2005;85:3-16.

62   Thalbourne MA, Bassett DL. *The Manic Depressiveness Scale: a preliminary effort at replication and extension.* Psychol Rep 1998;83:75-80.

63   Joseph S, Lewis CA. *The Depression-Happiness Scale: reliability and validity of a bipolar self-report scale.* J Clin Psychol 1998;54:537-44.

64   Bauer MS, Critis-Christoph P, Ball WA, et al. *Independent assessment of manic and depressive symptoms by self-rating.* Arch Gen Psychiatry 1991;48:807-12.

65   Folstein MF, Luria R. *Reliability, validity, and clinical application of the Visual Analogue Mood Scale.* Psychol Med 1973;3:479-86.

66   Giglio LM, Magalhães PV, Andreazza AC, et al. *Development and use of a biological rhythm interview.* J Affect Disord 2009;118:161-5.

67   Bowden CL, Singh V, Thompson P, et al. *Development of the bipolar inventory of symptoms scale.* Acta Psychiatr Scand 2007;116:189-94.

68   Dennehy EB, Suppes T, Crismon ML, et al. *Development of the Brief Bipolar Disorder Symptom Scale for patients with bipolar disorder.* Psychiatry Res 2004;127:137-45.

69   Wong G, Lam D. *The development and validation of the coping inventory for prodromes of mania.* J Affect Disord 1999;53:57-65.

70   Placidi GF, Signoretta S, Liguori A, et al. *The semi-structured affective temperament interview (TEMPS-I). Reliability and psychometric properties in 1010 14-26-year old students.* J Affect Disord 1998;47:1-10.

71   Altman EG, Hedeker DR, Janicak PG, et al. *The Clinician-Administered Rating Scale for Mania (CARS-M): development, reliability, and validity.* Biol Psychiatry 1994;36:124-34.

72   Brierley CE, Szabadi E, Rix KJ, et al. *The Manchester Nurse Rating Scales for the daily simultaneous assessment of depressive and manic ward behaviours.* J Affect Disord 1988;15:45-54.

73   Eckblad M, Chapman LJ. *Development and validation of a scale for hypomanic personality.* J Abnorm Psychol 1986;95:214-22.

74   Reilly-Harrington NA, DeBonis D, Leon AC, et al. *The interactive computer interview for mania.* Bipolar Disord 2010;12:521-7.

75   Krüger S, Quilty L, Bagby M, et al. *The Observer-Rated Scale for Mania (ORSM): development, psychometric properties and utility.* J Affect Disord 2010;122:179-83.