

## Review Article

# The International Society for Bipolar Disorders (ISBD) Task Force report on the nomenclature of course and outcome in bipolar disorders

Tohen M, Frank E, Bowden CL, Colom F, Ghaemi SN, Yatham LN, Malhi GS, Calabrese JR, Nolen WA, Vieta E, Kapczinski F, Goodwin GM, Suppes T, Sachs GS, Chengappa KNR, Grunze H, Mitchell PB, Kanba S, Berk M. The International Society for Bipolar Disorders (ISBD) Task Force report on the nomenclature of course and outcome in bipolar disorders.

Bipolar Disord 2009; 11: 453–473. © 2009 The Authors.  
Journal compilation © 2009 John Wiley & Sons A/S.

**Objectives:** Via an international panel of experts, this paper attempts to document, review, interpret, and propose operational definitions used to describe the course of bipolar disorders for worldwide use, and to disseminate consensus opinion, supported by the existing literature, in order to better predict course and treatment outcomes.

**Methods:** Under the auspices of the International Society for Bipolar Disorders, a task force was convened to examine, report, discuss, and integrate findings from the scientific literature related to observational and clinical trial studies in order to reach consensus and propose terminology describing course and outcome in bipolar disorders.

**Results:** Consensus opinion was reached regarding the definition of nine terms (response, remission, recovery, relapse, recurrence, subsyndromal states, predominant polarity, switch, and functional outcome) commonly used to describe course and outcomes in bipolar disorders. Further studies are needed to validate the proposed definitions.

**Conclusion:** Determination and dissemination of a consensus nomenclature serve as the first step toward producing a validated and standardized system to define course and outcome in bipolar disorders in order to identify predictors of outcome and effects of treatment. The task force acknowledges that there is limited validity to the proposed terms, as for the most part they represent a consensus opinion. These definitions need to be validated in existing databases and in future studies, and the primary goals of the task force are to stimulate research on the validity of proposed concepts and further standardize the technical nomenclature.

Mauricio Tohen<sup>a,b</sup>, Ellen Frank<sup>c</sup>, Charles L Bowden<sup>a</sup>, Francesc Colom<sup>d</sup>, S Nassir Ghaemi<sup>e</sup>, Lakshmi N Yatham<sup>f</sup>, Gin S Malhi<sup>g</sup>, Joseph R Calabrese<sup>h</sup>, Willem A Nolen<sup>i</sup>, Eduard Vieta<sup>d</sup>, Flávio Kapczinski<sup>j</sup>, Guy M Goodwin<sup>k</sup>, Trisha Suppes<sup>l</sup>, Gary S Sachs<sup>m</sup>, KN Roy Chengappa<sup>c</sup>, Heinz Grunze<sup>n</sup>, Philip B Mitchell<sup>o</sup>, Shigenobu Kanba<sup>p</sup> and Michael Berk<sup>q</sup>

Key words: affective switch – bipolar disorder – course and outcome – functional outcome – ISBD – nomenclature – predominant polarity – recovery – recurrence – relapse – remission – response

Received 15 October 2008, revised and accepted for publication 15 February 2009

Corresponding author:  
Mauricio Tohen, MD, DrPH, MBA  
Division of Mood and Anxiety Disorders  
Department of Psychiatry  
University of Texas  
Health Science Center at San Antonio  
7703 Floyd Curl Drive, MC 7792  
San Antonio, TX 78229, USA  
Fax: 210-567-3759  
e-mail: tohen@uthscsa.edu

## Introduction

The importance of developing a compendium of broadly accepted definitions of terms used to

characterize course and outcome in psychiatric disorders has been the focus of recent publications (1–5). Commonly used and understood terminology is essential to making meaningful comparisons across studies. For observational studies, common terminology would improve our ability to identify predictors of a variety of outcomes across patient

Affiliations and disclosure information for all authors are listed before the references.

populations worldwide. In clinical trials, the consistent use of standard terms to describe and compare meaningful outcomes (beyond statistical analysis of baseline-to-endpoint changes) using different therapies would ultimately be in the best interest of patients. The determination and validation of the proposed terminology should be based in clinically meaningful outcomes that can be applied to improve patient care.

In 1989, the MacArthur Foundation sponsored a set of meetings with the goal of achieving consensus in defining outcomes for clinical studies of unipolar depression. The result was a 1991 consensus report published by Frank et al. (5) which articulated the key elements inherent in applying the concepts of response, remission, recovery, relapse, and recurrence to unipolar major depressive disorder (MDD). The report recommended that these definitions be based on observable phenomena and include a temporal focus reflecting symptom change over the patient's lifetime. The proposed definitions, with the exception of the definition of response, did not imply a specific cause of symptom change, since symptom change can be a result of specific treatment effects, nonspecific effects of treatment, or the natural waxing and waning of depressive symptoms. More recently, an American College of Neuropsychopharmacology (ACNP) task force updated these original definitions, focusing on response and remission criteria in unipolar MDD (1), making relatively few changes from the original document, at least in part because data analyses to support or invalidate the original definitions were still lacking.

In 1994, the longitudinal course specifiers were introduced into the Diagnostic and Statistical Manual of Mental Disorders 4<sup>th</sup> Edition (DSM-IV) to describe course and outcome in mood disorders (6). The goal was to enable clinicians to describe the interepisode status of mood disorders. However, an important limitation of the DSM-IV course specifiers is the lack of quantitative criteria by which to determine recovery, leaving it to be based solely on clinical judgment (6).

Since the late 1980s, there has been a virtual explosion of research on the course and treatment of bipolar disorder; thus, it seems appropriate to return to the challenge of defining outcomes of this potentially devastating disorder.

Although operational definitions have been proposed to describe outcomes in bipolar disorders (7–9), the criteria used to define terms such as recurrence, relapse, response, remission, and recovery have varied both in observational studies and in clinical trials. Recently, Hirschfeld et al. (8) and Martínez-Arán et al. (9) proposed definitions for

some of the terms utilized in clinical trials in bipolar disorders.

Under the auspices of The International Society for Bipolar Disorders (ISBD), a task force was created with the purpose of establishing a consensus nomenclature on the course and outcome of bipolar disorders (2, 10). The first step was to assemble a group of 19 global experts in the field who, based on consensus, would compile definitions of nine terms commonly used in observational studies and clinical trials in bipolar disorder: response, remission, recovery, relapse, recurrence, subsyndromal states, predominant polarity, switch, and functional outcome. The task force was divided into seven sections, with each section addressing a particular concept or concepts: (i) response/remission/recovery for depression, (ii) response/remission/recovery for mania, (iii) overall relapse/recurrence, (iv) subsyndromal states, (v) predominant polarity, (vi) switch, and (vii) functional outcomes. Each section reviewed and discussed the published empirical evidence and made recommendations for each term/concept; once consensus was reached, it was reviewed by the whole task force. The task force had several face-to-face meetings and teleconferences wherein goals and consensus on the terminology were established. For each term, specific definitions have been summarized in Tables 1–7 and depicted in Figs. 1 and 2. Examples of proposed validation methodology have also been included. Each section presents a rationale and discussion of how consensus was reached, followed by a summary of the proposed term definition.

This first step needs to be followed by further field examination of the validity and ease of use of this terminology, as well as further development of validation methodologies. Validation should be carried out in existing databases, new observational studies, and new clinical trials. By developing a nomenclature that is validated, diagnostic classification systems such as the DSM-V and the International Classification of Diseases (ICD) may consider its adoption. Most importantly, having a validated, widely utilized nomenclature will allow investigators to compare information in areas of interest such as predictors of outcome with biomarkers or evaluation of new therapeutic agents. Finally, all terms need further validation to determine their clinical and public health relevance. Unfortunately, methodologies for the validation of terms on course and outcome in bipolar disorders have not yet been developed. Therefore, our recommendations should be considered only as a step to further develop a clinically meaningful terminology that is helpful to

individual clinicians, clinical epidemiologists, and public health experts.

### Response

Rationale and discussion of proposed terminology

Within the existent literature, response generally implies a clinically meaningful degree of symptom reduction during the course of a treatment intervention that is usually accompanied by an improvement in the patient's perception of the quality of his/her mood, daily functioning, and/or pain/distress. Identification of response is clearly useful to patients and clinicians, who must decide ultimately whether to continue, adjust the dose of, add to, or discontinue current treatment. These clinical decisions are inherently categorical and legitimately call for an outcome that provides a yes/no answer for each patient. The concept of response is temporally linked to onset or change during treatment (even if only watchful waiting), even though response, however defined, does not imply a causal relationship to the treatment itself.

There are, however, a number of limitations to using response as a predefined goal of treatment or as a primary outcome criterion in clinical trials. Response strongly depends on initial pretreatment symptom severity, and its ascertainment requires the systematic assessment of symptoms before and during treatment. Any unreliability in assessing initial symptom severity therefore directly affects the reliability of recognizing a response. Of note, when using symptom severity as a clinical trial inclusion criterion, if the same scale is used to measure symptoms at baseline and to determine entry into the study (above a certain threshold on a symptom rating scale), regression to the mean will further contribute to the invalid impression of symptomatic improvement (11). Furthermore, the recognition of a 'clinically significant' benefit depends on the initial state from which change is measured, the clinical purpose in ascribing response, and the clinical context. For example, a modest benefit in highly treatment-resistant bipolar depression may be more clinically significant than a greater benefit in a nonresistant depression. Specifically, while a  $\geq 50\%$  reduction in baseline severity is a commonly accepted convention, it may be inadequate for defining clinically significant benefit in a more severely ill or highly treatment-resistant patient (12), who may still be worse off at treatment exit (in terms of symptoms, behavior, functioning, or pain/distress) compared to another patient who does not respond but began with a less severe baseline depression (12). This may be

especially true in short-term treatment trials for bipolar depression. We recognize that functional improvement should be the goal of treatment. However, as stated in a recent issue of *Bipolar Disorders* addressing the topic of functional outcomes (13), the relationship between symptomatic and functional improvement is not fully understood. The task force felt it best at this point to exclude functional change as part of the definition of response and instead consider functional outcome as a separate parameter to be addressed further in this article. Finally, there is currently no clear recommended methodology on how to validate the term *response*.

Response in bipolar depression

*Symptomatic response—bipolar depression.* For unipolar depression, response has typically been defined as a  $\geq 50\%$  reduction in pretreatment symptom severity using symptom rating scales such as the Hamilton Rating Scale for Depression (HAM-D) (14), the Montgomery-Åsberg Depression Rating Scale (MADRS) (15), the Inventory for Depression Symptomatology (IDS) (16), and the Bipolar Depression Rating Scale (BDRS) (17). However, in the case of depression associated with bipolar disorder, we want to be especially mindful of whether a reduction in depressive symptoms is associated with a concomitant increase in symptoms of mania-hypomania. Therefore, to determine if a patient has responded successfully to a treatment for bipolar depression, he or she should demonstrate, by use of a standard mania rating scale such as the Young Mania Rating Scale (YMRS) (18) or the Clinical Global Impression for Bipolar Illness (CGI-BP) (19), no worsening in symptoms of mania, severity of mania (e.g., at least moderately ill), or change of mania (much worse or very much worse), which were determined to be in the normal (nonclinical) range at the outset of depression treatment. It would not be appropriate to specify a percent minimum worsening in mania scores, as patients may have very low scores; an alternative is to establish a maximum absolute score using any of the previously mentioned scales.

As mentioned above, the definition of response should represent a clinically meaningful benefit in the context of the population under study, taking into account treatment resistance, initial symptom severity, frequency of cycling, and other clinical factors. Of note, a recent report by Leucht et al. (4), exploring different percents of improvement to define response with the use of the Brief Psychiatric Rating Scale (BPRS) (20) in patients with schizophrenia, recommended a comprehensive approach.

Leucht and colleagues suggest replacing the usual one or several arbitrarily chosen cutoffs with incremental steps of 25% (< 25% reduction from baseline; 25–49% reduction; 50–74% reduction; 75–100% reduction). This approach, which has not previously been widely utilized in bipolar disorder studies, provides a whole range of responder rates that can be compared across studies.

To be certain that we are not simply observing random fluctuation in mood symptoms (even more common in patients with bipolar disorder than in those with unipolar disorder), we should consider requiring that response criteria be met for a specified period of time, typically 2–4 consecutive weeks, to take into account error in the assessment of symptomatology and any unstable symptomatic fluctuations. The task force consensus is to consider ascribing a *provisional* response when the response criterion is first met, then amending to a *definite* response when the response criterion is still met at the end of 2–4 weeks. The issue of validation of the term ‘response’ requires further consideration; however, to date no clear criteria for validation have been established. Validation criteria for response to be considered could include the ability to predict remission for a specified period of time, roughly 6 or 12 months. We recognize, however, that the ability to predict remission in order to validate response may be too simplistic for complex phasic conditions such as bipolar disorders.

*Syndromal response—bipolar depression.* The measurement of syndrome rather than symptom improvement has been suggested both for depression (5, 7) and for mania (7). The advantage of syndromal measurement improvement is its inherent association to the diagnostic classification being utilized. On the other hand, if the definition of syndromes changes, the use of existing symptom rating scales may not be appropriate for the new syndrome definition. The task force’s consensus is that currently both symptomatic and syndromal assessments should be considered.

For bipolar depression, the task force also recommends defining syndromal response only on the basis of improvement in the nine criterion symptoms of major depressive episode in the absence of symptoms of mania or hypomania, as specified in the DSM-IV, Revised Text (DSM-IV-TR) (6), or subsequently in the criterion symptoms enumerated in the DSM-V. Each criterion can be evaluated using a severity scale such as the CGI (range 1–7) and only symptoms with a minimum score (we suggest 4) at baseline can be evaluated. We recognize that these nine criteria have not been

widely used in a definition of response, as clinical trials have traditionally utilized symptom rating scales, but as mentioned above, the task force recommends both syndromal and symptom assessment. Furthermore, although we acknowledge that bipolar depressive episodes are largely comorbid, the task force recommends that noncriterion symptoms that are commonly associated with a major depressive episode, such as anxiety, panic attacks, irritability, hopelessness, avoidance, or cognitive dysfunction, not be included in the definition of syndromal response. Moreover, because these associated symptoms may be a function of other commonly concurrent Axis I, II, or III conditions (21), they may or may not respond to the treatment under study for bipolar depression. The issue of response of comorbid conditions needs further study. Regarding functioning, additional studies are needed to better define the norms for linking different levels of symptom reduction with different degrees of functional improvement. Since these associations are imperfect, it is important to know whether discrepancies in the degrees of symptom improvement and functional improvement have prognostic relevance in general, or are specific to particular subgroups of bipolar patients [e.g., bipolar I (BPI) versus bipolar II (BPII) disorder, or normal cycling versus rapid cycling].

#### Summary of task force recommendations (see Table 1)

The task force recommends further study of alternatives to current definitions of percent improvement and minimum duration, and recommends excluding social and occupational functioning in the definition of response. For *symptomatic response*, in addition to the commonly used 50% improvement, consideration should be given to cutoffs with incremental steps of 25% (< 25% reduction from baseline; 25–49% reduction; 50–74% reduction; 75–100% reduction). For *syndromal response*, the task force recommends a  $\geq 50\%$  improvement on each of the core symptoms of depression as defined in the DSM criteria that were present at baseline, without inclusion of comorbid symptoms.

Also to be considered is further developing clinically relevant proposed validation methodology, such as the ability to predict remission over a predetermined period.

#### Response in mania

*Rationale and discussion of proposed terminology.* The concepts of response, recovery, and remission have been extensively reported in the literature in

Table 1. Response

**Bipolar depression**

Rationale	<ul style="list-style-type: none"> <li>• Noncriterion symptoms (anxiety, panic attacks, irritability, hopelessness, avoidance, cognitive dysfunction) should not be included, as they may be a function of concurrent Axis I, II, or III conditions and may not respond to the treatment for bipolar depression</li> <li>• Response criteria should be met for a specified period of time to reduce error in assessment or observation of random fluctuation of mood symptoms</li> </ul>
Task force recommendation	<p><u>Syndromal</u></p> <ul style="list-style-type: none"> <li>• <math>\geq 50\%</math> improvement on each of the core symptoms of depression as defined in the DSM criteria assessing only those with scores of <math>\geq 4</math> as measured on a scale of 1–7</li> </ul> <p><u>Symptomatic</u></p> <ul style="list-style-type: none"> <li>• Percentage improvement using HAMD, MADRS, IDS, or BDRS (reduction from baseline): <math>&lt; 25\%</math>; 25–49%; 50–74%; 75–100%</li> <li>• Exclude social and occupational functioning</li> <li>• Ascribe <i>provisional</i> response when the response criterion is first met; amend to <i>definite</i> response when criterion is still met at the end of 2–4 consecutive weeks</li> <li>• Ability to predict remission over a subsequent predetermined period</li> </ul>
Validation methodology	

**Bipolar mania**

Rationale	<ul style="list-style-type: none"> <li>• Successful response to treatment for mania includes demonstrated lack of exacerbation of depressive symptoms</li> <li>• Response criteria should be met for a specified period of time to reduce error in assessment or observation of random fluctuation of mood symptoms</li> </ul>
Task force recommendation	<p><u>Syndromal</u></p> <ul style="list-style-type: none"> <li>• <math>\geq 50\%</math> improvement on each of the core symptoms of mania as defined in the DSM criteria assessing only those with scores of <math>\geq 4</math> as measured on a scale of 1–7</li> </ul> <p><u>Symptomatic</u></p> <ul style="list-style-type: none"> <li>• 50% improvement in mania symptom severity using YMRS or MRS</li> <li>• Percentage improvement (reduction from baseline): <math>&lt; 25\%</math>; 25–49%; 50–74%; 75–100%</li> <li>• Ascribe <i>provisional</i> response when the response criterion is first met; amend to <i>definite</i> response when criterion is still met at the end of 2–4 consecutive weeks</li> <li>• Ability to predict remission over a subsequent predetermined period</li> </ul>
Validation methodology	

HAMD = Hamilton Rating Scale for Depression; MADRS = Montgomery-Åsberg Depression Rating Scale; IDS = Inventory for Depression Symptomatology; BDRS = Bipolar Depression Rating Scale.

unipolar depression, schizophrenia, and (to some extent) in bipolar disorder. In the case of mania, it is important that these concepts evolve in a similar fashion as those for bipolar depression, including the concept that a reduction of manic symptoms should not be associated with a concomitant increase in depressive symptoms.

As in bipolar depression, the definition of response in mania should encompass a clinically meaningful benefit in the context of the population under study. For *symptomatic improvement* in mania, a symptom severity rating scale such as the YMRS or the Mania Rating Scale (MRS) (22) should be considered, as well as the use of incremental steps for symptom improvement ( $< 25\%$ ; 25–49%; 50–74%; 75–100%) from baseline. For *syndromic response*, we recommend a  $\geq 50\%$  improvement on each of the core symptoms of mania as defined in the DSM criteria. Also, as in bipolar depression, a  $\geq 50\%$  improvement on each of the core symptoms of mania as defined in the DSM criteria can be measured, using a severity scale such as the CGI (range 1–7) and only evaluating symptoms with a minimum score (we suggest 4) at baseline.

In the context of mania, we must also consider any concomitant increase in depression. For a patient to be considered as having responded successfully to a treatment for mania, he or she should demonstrate lack of exacerbation of depressive symptoms, as defined by not exceeding an absolute score such as on the HAMD, MADRS, IDS, or BDRS, or a significant change in the CGI-BP severity of depression or change in depression.

To be certain that we are not simply observing random fluctuation in mood symptoms, we suggest using the same time period (2–4 weeks) as for bipolar depression. Furthermore, we also suggest, for both bipolar depression and mania, consideration of the concept of *provisional* response (when a  $\geq 50\%$  improvement is first met), followed by *definite* response (when the response criteria have been met for 2–4 weeks).

Summary of task force recommendations (see Table 1)

Recommendations are similar to those for bipolar depression, including the use of step increments of 25%, absence of exacerbation of the opposite pole, and use of the concepts of provisional and definite

response. For syndromal response, we recommend utilizing only the core DSM symptoms. In addition, as in depression, clinically relevant proposed validation methodology is needed, such as the ability to predict remission over a predefined period.

## Remission

Concept of remission

Remission implies that the signs and symptoms of a specified condition (e.g., depression) are absent or nearly absent, and that there has not been a concomitant increase in symptoms of another condition (e.g., mania or hypomania). Remission, unlike response, entails an absolute allowable ceiling level in symptom expression and the (near) absence of symptoms.

The advantage of using remission criteria is that it estimates the number of subjects having no significant symptoms at the end of a study. This is in contrast to response, which is defined in such a way that subjects who started a study with a YMRS score of 50 (1–60 scale) or a MADRS score of 40 would be still be clearly symptomatic, despite having achieved a 50% reduction in baseline severity. A limitation of the concept of remission is that it does not reflect the change of symptoms. In a study in which many participants have low symptom scores at baseline, many may be in remission at study completion even though the overall decrease of symptoms is small. Therefore, the use of both remission and response criteria may add value to particular studies.

Remission in bipolar depression: rationale and discussion of proposed terminology

*Syndromal remission.* We recommend that syndromal remission refer only to the nine criterion symptom domains identified in DSM-IV-TR to diagnose a major depressive episode (MDE). Should the definition of MDE change, e.g., core criterion symptoms added or deleted, operationalizing remission will require revised methods. This recommendation is consistent with the previous recommendation for response to use solely the DSM criterion symptoms, and is based on the evidence to date that demonstrates the relevance of remission to functioning and prognosis (i.e., most studies have focused on core depressive symptoms). For each criterion symptom, a severity score could be added based on a scale such as the CGI (range 1–7). A definition can be operationalized to allow the presence of minimum symptomatology

(e.g., fewer than three core symptoms with a score of 3 or more within a range of 1–7) instead of complete absence of symptoms.

Noncriterion symptoms or associated features may be of use as secondary outcomes, although there are insufficient data to date on this issue.

We recommend that daily functioning should not be part of the definition of remission for the same reasons noted for response. In bipolar disorder, symptomatic remission is not necessarily associated with a return to premorbid day-to-day functioning. Thus, functioning should be measured and reported as a separate outcome.

We recommend that neither sad mood nor loss of interest/pleasure should be present in the remitted state and that fewer than 3 of the 7 remaining core criterion symptoms (e.g., poor concentration, disturbed appetite/weight, disturbed sleep, etc.) would be meaningfully (score of  $\geq 3$  within a range of 1–7) present. We recognize that the highly specific nature of this recommendation deserves study despite its face validity. We felt that the presence of either essential symptom (loss of interest/pleasure, sad mood) would likely be associated with a worse prognosis than if both were absent, and that a simple count of symptoms (e.g., presence of 3 or 4 as opposed to 5 of the 9 criterion symptoms) provided an incomplete description of the remitted state. The basic notion underlying this recommendation is that depression at its core represents a hedonic deficit that is best captured by these two depressive symptoms. Thus, if either symptom were present, the disorder would not be truly remitted. In addition, the CGI-BP severity of depression and severity of mania scores should both be  $\leq 2$ .

The proposed definition of syndromal remission logically requires that any assessment used to operationalize remission must include the entire criterion symptom domains used to diagnose an MDE by DSM-IV-TR or subsequently DSM-V. Given the definition of remission above, the ideal way to determine remission is via a structured clinical interview that ascertains the criterion symptoms of depression and mania. This can be accomplished by repeating the depression and mania sections of the Structured Clinical Interview for DSM-IV (SCID). Abridged versions of the SCID, such as the Mini International Neuropsychiatric Interview (MINI) (23), can also be used.

*Symptomatic remission.* Currently, most clinicians estimate remission using total score thresholds, most often utilizing either the HAMD, the BDRS, or the MADRS, without reference to the above-recommended syndromal definition. Rating scales that identify all nine criterion domains include the

nine-item self-reported Patient Health Questionnaire (PHQ-9) (24), the 16-item Quick Inventory of Depressive Symptomatology, available as a clinician rating (QIDS-C<sub>16</sub>) or self-report (QIDS-SI<sub>R</sub>), (25, 26), and the Beck Depression Inventory, Version II (BDI-II) (27, 28), a self-report. The BDI-II does not include weight gain, but does otherwise include all other criterion symptom domains. A concurrent determination of the absence of mania/hypomania must also be made using a scale such as the YMRS or the MRS.

If one chooses the HAMD-17 to estimate remission, we recommend that a score  $\leq 5$  or  $\leq 7$  (based on the precedent in the literature) be used. It is noteworthy that in looking at studies of unipolar disorder, Nierenberg et al. (29) found that only 17.6% of patients with a HAMD-17 score  $\leq 7$  had no symptoms of MDD. A HAMD-17 score  $\leq 7$  corresponds to a MADRS score  $\leq 9$  (30) or a 30-item IDS-Clinician-Rated (IDS-C<sub>30</sub>) score  $\leq 12$ , an IDS Self-Report (IDS-SI<sub>R</sub>) score  $\leq 14$  (30), or a QIDS-C<sub>16</sub> or QIDS-SI<sub>R</sub> score  $\leq 5$  (25, 26). The corresponding PHQ-9 score is likely  $\leq 5$  (24). Alternatively, Zimmerman et al. (31–33) have recommended a MADRS total score  $\leq 5$  to define remission. Berk et al. (34) found that scores  $< 5$  in the MADRS scale correlate better with a CGI score = 1.

However, it is important to note that total scores on selected rating scales are insufficient (and thus not recommended to be used alone) to ascertain remission if one uses our proposed definition that rests on the nine core criterion symptoms. For example, the HAMD-17 does not include oversleeping, weight and appetite increase, or impaired concentration/decision making. The MADRS (15) does not include oversleeping, overeating, loss of interest (although it assesses inability to feel), loss of energy (although it assesses lassitude), self-criticism (guilt), or psychomotor changes.

The newly developed BDRS may have an advantage over previous scales, as it includes oversleeping, overeating, self-reported depressed mood, reduced social engagement, impaired concentration, agitation, and guilt. For the BDRS, the authors recommend a score for remission  $\leq 8$ .

Summary of task force recommendations (see Table 2)

We recognize that the above recommendations in which remission is defined in terms of ‘minimal’ symptoms are somewhat arbitrary. These and alternative definitions call for empirical studies that relate different symptom and duration criteria to prognosis. Unlike for unipolar depression, we do not recommend utilizing a duration criterion,

Table 2. Remission

Rationale	<ul style="list-style-type: none"> <li>• Symptomatic remission may not be associated with a return to premorbid day-to-day functioning, which should be measured and reported as a separate outcome</li> <li>• For bipolar depression, presence of sad mood and/or loss of interest/pleasure may be associated with a worse prognosis, and a simple count of symptoms (e.g., presence of 3 or 4 versus 5 of the 9 criterion symptoms) incompletely describes remission</li> <li>• No duration criterion</li> </ul>
<b>Bipolar depression</b>	
Task force recommendation	<p><u>Syndromal</u></p> <ul style="list-style-type: none"> <li>• Sad mood and/or loss of interest/pleasure may not be present, and <math>&lt; 3</math> of the 7 remaining core criteria may be meaningfully (score <math>&gt; 3</math> within a range of 1–7) present</li> <li>• CGI score <math>\leq 2</math></li> </ul> <p><u>Symptomatic</u></p> <ul style="list-style-type: none"> <li>• HAMD-17 score <math>\leq 5</math> or <math>\leq 7</math></li> <li>• MADRS score <math>\leq 5</math> or <math>\leq 7</math>, or BDRS score <math>\leq 8</math></li> <li>• Exclude daily functioning</li> <li>• Ability to predict recovery over a predetermined period</li> </ul>
Validation methodology	
<b>Bipolar mania</b>	
Task force recommendation	<p><u>Syndromal</u></p> <ul style="list-style-type: none"> <li>• Focused on core affective symptoms</li> <li>• Focused on the 7 criterion symptom domains identified in DSM to diagnose a manic episode</li> <li>• DSM-IV criterion A <math>\leq 2</math>; no B criterion rated <math>&gt; 3</math>; no more than two B criteria = 3</li> <li>• Initial mixed episodes fulfill recovery criteria; no depression criterion <math>&gt; 3</math>; no more than three criteria = 3</li> <li>• CGI-BP <math>\leq 2</math></li> </ul> <p><u>Symptomatic</u></p> <ul style="list-style-type: none"> <li>• YMRS <math>&lt; 8</math> or <math>&lt; 5</math></li> <li>• Ability to predict recovery over a predetermined period</li> </ul>
Validation methodology	

HAMD = Hamilton Rating Scale for Depression; MADRS = Montgomery-Åsberg Depression Rating Scale; BDRS = Bipolar Depression Rating Scale; CGI-BP = Clinical Global Impression for Bipolar Illness; YMRS = Young Mania Rating Scale.

but rather that a duration criterion be specified for the term ‘recovery’.

The second issue is the imperative to develop rating scales that adequately assess some of the nuances of bipolar depression; the BDRS or the IDS, recently administered in clinical trials (35, 36), and the revision of the HAMD by Thase et al. (37) are examples of attempts to achieve this goal.

The third issue is to develop clinically relevant proposed validation methodology for remission, such as the ability to predict recovery over a predetermined period.

Remission in mania: rationale and discussion of proposed terminology

*Syndromal remission.* For syndromal remission, we recommend following a similar concept as for depression and focusing on only the seven criterion symptom domains identified in DSM-IV-TR and subsequently in DSM-V to diagnose a manic episode. For DSM-IV, we recommend consideration of operationalized criteria such as utilized by Tohen et al. (7), which are parallel to those recommended for depression by Frank et al. (5) and most recently by Rush et al. (1). For remission of mania, the following criteria need to be met: DSM-IV criterion A for mania  $\leq 2$  (range 1–7), with no B criterion rated  $> 3$  and no two B criteria rated = 3. Patients with initial mixed episodes need to fulfill recovery criteria for both a manic and a depressive episode (no depression criterion  $> 3$  and no more than three criteria = 3). In addition, CGI-BP severity of mania and severity of depression ratings are both to be  $\leq 2$ .

*Symptomatic remission.* In most recent studies, symptomatic remission has been measured with the YMRS or the MRS. Although different scores have been used, perhaps the most commonly applied has been a score  $< 12$ . However, more conservative scores have also been applied. In an observational study, Tohen et al. (7) defined remission as  $< 5$ ; Berk et al. (34) recently reported that a YMRS score  $< 4$  approximates a CGI-BP score of 1; and Chengappa et al. (38) showed that a YMRS score  $< 8$  is equivalent to a patient’s ability to function (minimum symptomatology).

Summary of task force recommendations (see Table 2)

We recognize that, for symptomatic remission, the most commonly applied score utilizing the YMRS has been  $< 12$ , but we are of the opinion that a more conservative score such as  $< 8$  or  $< 5$

should be utilized. For syndromal remission, we recommend focusing on core affective symptoms. As in bipolar depression, we do not recommend utilizing a duration criterion. We recognize that, as in bipolar depression, the above recommendations are somewhat arbitrary. Therefore, these and alternative definitions call for empirical studies that relate different symptom and duration criteria to prognosis. In addition to conducting additional observational studies, existing clinical trial databases can be utilized to further validate these definitions. As in depression, clinically relevant proposed validation methodology for remission of mania needs to be tested, such as the ability to predict recovery over a predetermined period, e.g., 6 or 12 months.

## Recovery

Rationale and discussion of proposed terminology

The concept of recovery implies an extended period of remission such that neither an MDE nor an episode of mania or hypomania is likely to occur in the near future. That is, recovery implies that the remitted state has persisted long enough and has sufficient consistency that many future months of remission can be anticipated for most patients. According to our conceptualization, recovery should only be ascribed after sufficient time has passed (e.g., 4–8 weeks in most patients) such that the recovered state is likely to persist for a reasonable period of time (e.g., 6–12 months in most patients). The National Institute of Mental Health Collaborative Study of Depression (39) and the McLean-Harvard First Episode project (7, 40) defined recovery as remission for at least eight weeks. The Systematic Treatment Enhancement Program for Bipolar Disorder project (41) defined recovery as two or fewer syndromal features of mania, hypomania, or depression, for at least eight weeks. Additional observational studies and clinical trials are needed to test the current assumption that a distinction between remission and recovery is indeed meaningful.

It seems clear that, in the majority of individuals with bipolar disorder, the underlying vulnerability to subsequent syndromal episodes remains indefinitely. Thus, in this set of definitions, ‘recovery’ does not refer to recovery from the illness, but from the last mood episode.

As is the case with remission, recovery may be ascribed while the patient is either on or off treatment. Recovery, once present, can only be lost if a recurrence occurs (i.e., subsyndromal manifestations of either depression or hypomania



may occur without loss of the ‘recovered’ status). Recovery may persist or be followed by a new episode (recurrence). When symptoms that are insufficient to meet the criteria for mania appear during or following recovery, the terms ‘subsyndromal mania’ or ‘hypomania following recovery’ are recommended.

We recommend that recovery should be defined only by symptomatic status for the same reasons that symptom status alone should be used to define remission. As in remission, recovery does not require normalization of day-to-day functioning, although such normalization may occur for some patients.

Recovery should be ascribed once a sufficient period of time has passed such that the recovered state is likely to persist for a reasonable period of time (6–12 months) in most patients. We recommend that eight consecutive weeks must pass, during which each week is characterized by meeting remission criteria for depressive and manic or hypomanic symptoms, before recovery can be ascribed. The task force estimates that one might expect recovery to be sustained if it is present for at least eight consecutive weeks (to ensure that transient fluctuations are not designated as recovery). Of note, DSM-IV uses the term ‘full remission’ when remission criteria have been met for eight weeks. Riso et al. (42) used a six-month duration period to define recovery in unipolar depression, with evidence of validation based on the prior course of illness. By definition, recovery can only occur after remission has been ascribed. The main reason for the eight-week recommendation is that placebo-controlled trials of continuation therapy and observational studies in patients following an episode of bipolar depression or mania indicate that the majority of new episodes occur within the first two months following the remission of the index episode (7, 39–41, 43). To ensure that recovery has occurred, measurements must be made frequently enough (i.e., every 1–2 weeks) to detect a return of the index episode.

We note that the term ‘recovery’ by patient advocacy and support groups often refers to subjective (personal) versus objective standards of recovery, based largely on social and occupational

functioning in contrast to the objective standards based on symptom severity and duration that we are proposing for use in clinical trials and observational studies. The task force adheres to the objective definition proposed here for those purposes, but is sensitive to the issues raised by patients who wish to frame their personal goals in a more positive way than our preferred terminology and objectives permit.

Summary of task force recommendations (see Table 3)

We recommend a minimum period of two months in remission as the definition of recovery. Further research to test the recommended definition is needed. Proposed clinically relevant validation to be considered includes the ability to predict absence of recurrence over a predetermined period, e.g., 12 months. Further consideration should also be given to the prediction of occupational and social recovery.

**Relapse and recurrence**

Rationale and discussion of proposed terminology

Definitions of relapse and recurrence depend on the clinical epidemiology of the condition being studied. In the case of bipolar disorders, mood episodes have different durations. There is a tendency for manic episodes to be briefer than mixed episodes, and for both to be briefer than depressive episodes. Moreover, the duration appears to be longer in patients with rare (e.g., once every 2–3 years) versus frequent (e.g., a rapid-cycling course) episodes. Thus, no uniform a priori definitions of relapse or recurrence in bipolar disorder can be given; rather, such definitions should be made with reference to the natural course of illness. Furthermore, such natural history data are best obtained in an untreated condition, since treatment is intended to reduce the duration of a current episode as well as the frequency of further episodes (thus altering the course of the illness), but such a research approach would be unethical.

Consensus definitions of relapse and recurrence have been proposed in unipolar depression, but so

Table 3. Recovery

Rationale	<ul style="list-style-type: none"> <li>• Recovery does not refer to recovery from the illness, but from the last mood episode</li> <li>• Recovery may be ascribed while on or off treatment. Once present, it can only be lost if recurrence occurs (not for subsyndromal manifestations of either depression or hypomania)</li> </ul>
Task force recommendation	<ul style="list-style-type: none"> <li>• <b>8 consecutive weeks</b> characterized by the virtual absence of depressive and manic or hypomanic symptoms</li> </ul>
Proposed validation methodology	<ul style="list-style-type: none"> <li>• Ability to predict absence of recurrence over a 12-month period</li> </ul>

far no such effort has been made in bipolar disorder. In unipolar depression, Frank et al. (5) suggested a consensus proposal for definitions of relapse as “a return of symptoms satisfying the full syndrome criteria for an episode that occurs during a period of remission, but before recovery” (i.e., sustained remission during  $\geq 8$  weeks) and of recurrence as “the appearance of a new episode of major depressive disorder, occurring during recovery.” In addition, they proposed operational criteria for remission (a HAMD-17 score  $\leq 7$  lasting  $\geq 2$  but  $< 8$  weeks) and for recovery (a HAMD-17 score  $\leq 7$  lasting  $\geq 8$  weeks). Thus, they not only discriminated remission and recovery, but also relapse as an early return of the syndrome (within eight weeks) and recurrence as a late return of the syndrome (after eight weeks). As described earlier in this article, we have taken a similar approach in differentiating remission and recovery.

The considerations for defining relapse and recurrence in unipolar depression appear to be consistent with the natural history literature that finds that most major depressive episodes in that condition last about 6–12 months. In bipolar disorder, the natural history literature (especially based on the pretreatment era) indicates that major depressive episodes last about 3–6 months. Pure manic episodes appear to last 2–4 months untreated and mixed episodes last about six months or longer. Ghaemi and others (44, 45) have suggested that perhaps in bipolar disorder, definitions of relapse or recurrence may need to be sensitive to mood phase subtype as well. Another issue raised by Ghaemi (44) is consideration of additional definitions in the case of a rapid-cycling course.

Currently used definitions

In recent studies addressing maintenance treatment of bipolar disorder, investigators have used various definitions for relapse or recurrence, and unfortunately have often used the terms interchangeably. Some have defined them according to diagnostic criteria for a manic, hypomanic, or major depressive episode [DSM-III-R criteria (46) or DSM-IV criteria (47)]; others according to a score on global rating scales [a four-point-scale using the morbidity index (0 = no symptoms to 3 = hospitalization) or a six-point scale applied by clinicians (1 = no disturbance to 6 = extremely severe recurrence) (48)]; or according to a score on a severity rating scale [MRS score  $\geq 16$  or Depressive Symptoms Scale score  $\geq 25$  (49), YMRS score  $\geq 15$  or HAMD score  $\geq 15$  (46, 50), or YMRS score  $\geq 20$  or 24-item HAMD score  $\geq 20$  (51)]; still others according to

the need for an intervention [with antidepressants (49), with additional medication or electroconvulsive therapy (52, 53), or with hospitalization (47)]; or to withdraw the patient from the study because of symptoms (49).

Task force proposal consideration

The overall consensus of the task force was that further study of definitions that take into account the polarity of the index episode is needed before applying different definitions depending on the nature of the previous episode. Empirical studies need to determine if the index episode determines the definitions of relapse and recurrence. For depressive episodes, relapse (i.e., the return of the index episode) would be defined as occurring up to eight weeks (i.e., two months) after remission from the acute episode, which may mean at 4–6 months after episode onset when treatment is started within two months after onset. Switch (i.e., the appearance of an episode of the opposite pole directly from/after the index episode) would be defined as occurring up to eight weeks after remission, and recurrence (either depression or mania) as occurring  $> 8$  weeks after remission from the acute episode (Fig. 1).

For manic episodes, relapse would be defined as occurring up to four weeks (i.e., one month) after recovery from the acute episode, which may mean 2–3 months after episode onset when treatment is started within 1–2 months after onset. Switch to depression would then also be defined as occurring up to four weeks after remission, and recurrence (either mania or depression) as occurring  $> 4$  weeks after remission from the acute episode (Fig. 2). A related consideration proposed in a leading text in our field (45) is to define recovery as

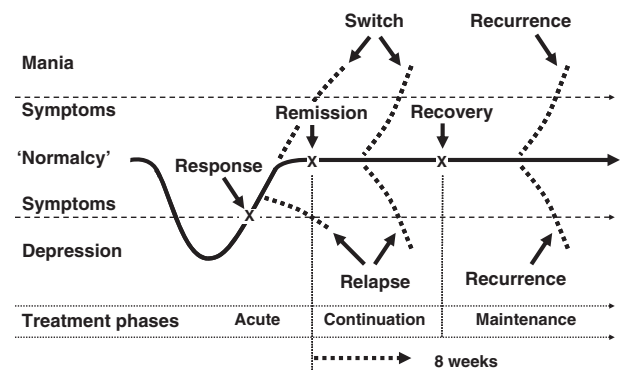


Fig. 1. Proposed definitions for remission, recovery, relapse, switch, and recurrence in index depressive episodes [taken from Frank et al. (5)].

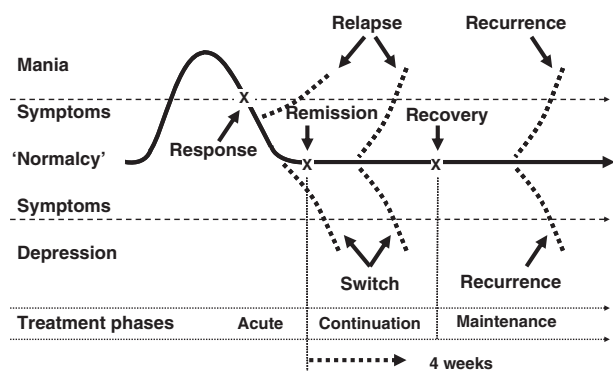


Fig. 2. Proposed definitions for remission, recovery, relapse, switch, and recurrence in index manic episodes [taken from Frank et al. (5)].

a duration of six months in remission rather than eight weeks, which suggests that the natural duration of most mood episodes ranges 2–6 months or even longer depending on the type of episode (shorter for mania, longer for depression, longest for mixed episodes). Others have suggested that mixed episodes are of intermediate duration (39). If we treat all mood episodes equally, and provide a single definition for all types, then the longer duration of six months would seem to be a rational cutoff point for recovery, after which episodes would represent new recurrences. Thus, an alternative to the eight-week criterion for recovery may be to explore longer duration, such as six months, which could prove to be a more valid definition.

Summary of task force recommendations (see Table 4)

Considering the paucity of empirical data supporting differentiation of the definition based on the previous episode, the task force currently recommends a single definition regardless of the index episode. Thus, relapse and recurrence would be defined as a new episode occurring within eight weeks and > 8 weeks, respectively, of remission of the index episode.

Table 4. Relapse and recurrence

Rationale	<ul style="list-style-type: none"> <li>• Durations of manic and depressive episodes differ, as do normal or rare versus rapid-cycling episodes. Definitions of relapse/recurrence should consider mood phase subtype and time course</li> <li>• There is a paucity of empirical data to base definitions on the previous episode</li> </ul>
Task force recommendation	<ul style="list-style-type: none"> <li>• A single definition regardless of index episode. Relapse and recurrence is defined as a new episode occurring within 8 weeks and &gt; 8 weeks, respectively, after having achieved remission from the index episode</li> </ul>
Validating methodology	<ul style="list-style-type: none"> <li>• To determine if a new episode before or after 8 weeks predicts time to remission/recovery or to further new episodes.</li> <li>• To determine if the type of index episode matters in the occurrence and type of new episodes</li> </ul>

Subsyndromal states

In both BPI and BPII patients, manic and depressive symptomatology that does not meet syndromal severity has consistently been reported as more common than syndromal manic or depressive states (54). Daily self-rated mood charting (a useful technique to study long-term illness expression in bipolar disorders) employed in a three-year study of BPI or BPII disorders indicated that patients experienced subsyndromal depressive symptoms approximately 25% of the time (54). Subsyndromal mood states are also associated with an increased risk of relapse compared to full syndromal mood episodes (7, 55–57), and studies also suggest that subsyndromal states are associated with impairments in cognition, functional ability, and health-related quality of life (QOL) (7, 55, 56, 58, 59). The scientific study of bipolar disorder, as well as clinical practice, would be benefited by operationalized definitional guidelines for subsyndromal symptomatology. In some ways, this aspect of bipolar symptomatology is defined by exclusion. Patients who have syndromal episodes currently subclassified as manic, mixed, hypomanic, or depressed would generally not be classified within a subsyndromal group. Similarly, patients who meet criteria for remission status would not be classified as subsyndromal. Both syndromal and remission threshold scores are discussed in other sections of this article. We approach subsyndromal definitions by first accepting that criteria for defining a patient as syndromal or remitted should be taken into account, and that subsyndromal severity scores should be continuous with those two boundaries.

Rationale and discussion of proposed terminology

The explicit operational definitions described here have some caveats. They have not been tested a priori in clinical trials, although they have been applied in some observational studies. They utilize the currently most frequently employed scales, which have major limitations. Most were devel-

oped in studies of hospitalized, full-episode patients (60). There is substantial reason to expect that they are less sensitive to mild symptomatology than to fully syndromal states (61, 62). The depression scales are further limited by not having been psychometrically developed utilizing solely, or even a majority of, bipolar patients. Neither the depression nor mania scales were developed with any patients who had BPII.

Several recently developed scales that do not have these limitations should provide greater sensitivity for the goals of assessing and studying subsyndromal states in relation to other illness-related variables (26, 63).

An additional limitation of the scales for which we provide threshold scores is that they have limited capability to test the full range of behavioral components of bipolar symptomatology. Factor analytic studies indicate principal components of hyperactive, impulsive, manic cognitive, depressive, anxious, psychotic, and irritable behavioral states in bipolar disorders (62, 64). In particular, the most frequently employed scales inadequately assess symptoms of anxiety, which are prominent in all phases of bipolar disorders (65). We therefore recommend that these guidelines be applied when other rating scales are utilized, and anticipate that the use of scales that cover the spectrum of symptom domains in bipolar disorders and have been developed around the full spectrum of bipolar disorder presentations, including subsyndromal and remitted patients, will allow finer-grained assessments. Recent work on the IDS and the corresponding QIDS, as well as the Bipolar Inventory of Symptoms Scale (BISS) (66) and the Structured Clinical Interview for Mood Spectrum, in the interview and self-report format (SCI-MOODS and MOODS-SR) would appear suitable for such use (64, 66, 67). Recent reports suggest that the SCI-MOODS/MOODS-SR also have good psychometric properties (67–69).

One further caveat to subsyndromal states is that we do not recommend operational guidelines based on self-rated mood charts, although we recognize the clinical utility and the secondary supportive hypotheses that they can test.

Hypomania occupies an unusual double role, being descriptive both of subsyndromal manic states in patients with BPI disorder and of the defining symptoms for diagnosis of BPII disorder. The establishment of syndromal criteria for BPII disorder, e.g., selection of specific symptoms, duration of symptoms, and any other illness course features, is beyond the scope or intent of this article.

#### Combining subsyndromal depression and mania scores

Operationally, any subject would fall outside the subsyndromal category within a study analysis on any rating period for which his/her score was in the syndromal range for either depression or mania. However, a patient would be required to have both depressive and manic scores in the remission range to fall outside the subsyndromal range for recovery. A subject with remission status on a depression scale who was within the subsyndromal range for mania would be categorically assessed as subsyndromal for mania. The same boundary scores for mania would apply regardless of whether a patient was classified as BPI or BPII.

We recommend guidelines to establish subsyndromal symptom states in bipolar disorders and to apply them in clinical studies by utilizing both established clinician-rated scales and other scales in development that may cover additional dimensions of bipolar symptomatology and provide more sensitive indicators of change. Subsyndromal status is a dimensional versus categorical approach and concept. The notion of a subsyndromal clinical status asserts that simply failing to meet syndromal criteria does not mean that the patient is well. Further, it assumes a lower symptom boundary beyond which the patient should be considered recovered. Viewing it as dimensional avoids stating which combination of symptoms, and their duration, constitutes a syndrome. Additionally, it avoids the conundrum of whether the symptoms are functionally impairing. Therefore, a patient with BPI or BPII disorder who has a total YMRS or MRS score of 10 is symptomatic and subsyndromal for the period under study, and cannot be considered as having ‘days well’ for the period. Clinically relevant validation methodology for subsyndromal symptoms to be considered includes the ability to predict relapse/recurrence over a 12-month period as recently reported by Judd et al. (70) and Tohen et al. (57). Other validation methodology to be considered includes the ability to predict functional impairment.

#### Summary of task force recommendations (see Table 5)

*Subsyndromal depression.* Based on syndromal definitions of depression in bipolar disorders, setting a score of around 15 as a threshold for moderately severe depression on all three of the HAMD, BDRS, and MADRS scales, we recommend as an upper boundary for subsyndromal depression a total score of 14 on the HAMD or MADRS and a score of 16 on the BDRS. Considering a score of 7 as an upper boundary

Table 5. Subsyndromal states

Rationale	<ul style="list-style-type: none"> <li>• Subsyndromal status is dimensional versus categorical. A subsyndromal clinical status that does not meet syndromal criteria cannot be interpreted as the patient being well. It also assumes a lower symptom boundary representing remission. Viewing subsyndromal states as dimensional relaxes the need to identify a syndrome based precisely on symptoms, their duration, and whether or not they are functionally impairing</li> </ul>
<b>Subsyndromal depression</b>	
Task force recommendation	<ul style="list-style-type: none"> <li>• Score of 14 on HAMD or MADRS and score of 16 on BDRS as upper boundary</li> <li>• Total score of 8 on HAMD or MADRS and score of 9 on BDRS as lower boundary</li> </ul>
<b>Subsyndromal mania</b>	
Task force recommendation	<ul style="list-style-type: none"> <li>• Score of 14 on MRS or YMRS as upper boundary</li> <li>• Total score of 8 on MRS or YMRS as lower boundary</li> </ul>
Validating methodology	<ul style="list-style-type: none"> <li>• Ability to predict relapse/recurrence over a 12-month period</li> </ul>

HAMD = Hamilton Rating Scale for Depression; MADRS = Montgomery-Åsberg Depression Rating Scale; BDRS = Bipolar Depression Rating Scale; MRS = Mania Rating Scale; YMRS = Young Mania Rating Scale.

for defining remission status with the HAMD or MADRS, and 8 for the BDRS, we recommend a total score of 8 as a lower boundary for subsyndromal depression on the HAMD or MADRS, and 9 for the BDRS.

*Subsyndromal mania.* Based on syndromal definitions of mania in bipolar disorders, setting a score of around 15 as a threshold for moderately severe mania on both the MRS and YMRS scales, the task force recommends a total score of 14 as an upper boundary for subsyndromal hypomania on the MRS or YMRS. Based on the use of a score of around 7 as an upper boundary for defining remission status with the MRS or YMRS, we recommend a total score of 8 as a lower boundary for subsyndromal mania on the MRS or YMRS.

The task force recommends that these upper and lower boundaries apply in studies of both BPI and BPII patients. If future studies on the course of illness of patients with BPII provided reasons to modify these boundaries, such amendments could then be undertaken. Validating criteria to be considered include the ability to predict relapse/recurrence over a predetermined period.

**Predominant polarity**

The predominant polarity of mood episodes of patients’ lifelong history may have critical relevance for their clinical management. However, there is currently no taxonomic translation for this clinically based concept.

Rationale and discussion of proposed terminology

Colom et al. (71) suggested a novel concept of putative therapeutic relevance based on the classification of patients with bipolar disorder as either predominantly depressed (PD) or predominantly manic (PM), as defined by having at least two-

thirds of their lifetime episodes at one polarity or the other. This proposed course specifier may be a valid prognostic parameter of particular relevance to long-term therapeutic decisions and prediction of outcome.

The available literature suggests that between 45% and 70% of all patients with bipolar disorder fulfill the suggested criteria for a predominant polarity (72–74), while the rest appear of undetermined polarity (PU). Among those with a defined polarity, about 50–60% of patients are PD and about 40% are PM (71–73). To our knowledge, all the studies performed using the concept of predominant polarity except one (74) show the same trend toward a majority of PD patients. If we consider all the patients, including those who are PU, depressive polarity would account for 25–35%. PD polarity is strongly associated with depressive onset, while PM polarity is associated with manic onset (71, 72, 75). PD polarity is more common among patients with BPII disorder (72, 73) and is associated with a higher number of years undiagnosed (73). Interestingly, if we do not use the concept of predominant polarity and only consider the number of days spent experiencing a certain episode, depression appears to be the predominant polarity in those with both BPI and BPII disorder (54, 76, 77).

Regarding the proposal for this new course specifier, a question arises relative to mixed episodes. Should we consider a new category such as ‘mixed polarity’ or should we relate mixed episodes to a certain polarity? At this point, not enough empirical information is available to support its inclusion. The link between first-episode polarity and subsequent polarity has been shown in several studies (71, 72, 75, 77), and might be a robust predictor of outcome when polarity of first episode is prospectively observed. Regarding validation, there is currently no standard methodology to validate the term. Clinically relevant

validation methodology of predominant mood polarity to be considered includes the ability to predict time to relapse/recurrence over a predetermined period of time, e.g., 12 months, and the ability to predict the polarity of the relapse/recurrence. In addition, investigators are urged to explore the validity of a predominately mixed group.

Summary of task force recommendations (see Table 6)

We recommend utilizing the definition proposed by Colom et al. (71) for predominant polarity, i.e., at least two-thirds of lifetime episodes being of one polarity or the other (PD or PM). Currently there is not enough empirical information to support the inclusion of a predominately mixed category.

**Treatment-emergent affective switch**

Rationale and discussion of proposed terminology

Like many areas in medicine where we confront a chronic condition characterized by an irregular course of acute symptoms, the problems of definition and attribution of ‘affective switch’ present interrelated challenges. Like attribution of treatment as a cause of improvement, the definition of affective switch is at best operational and leaves room for considerable error. The implication of causality cannot be settled as a categorical yes or no, but as more or less likely based on specified criteria. The definition should be chosen to maximize consistent application in research and provide predictive value for clinicians. Referring to these events as a ‘treatment-emergent affective switch’ (TEAS) seems preferable because it does not attribute causality. Terms like ‘antidepressant-induced switch’ or ‘antipsychotic-induced switch’ are error prone and hazardous.

The main criteria for TEAS are the amplitude and duration threshold for identifying a change (signal detection) and the time window from the putative causal intervention (last change in treatment). Use of episode criteria (based on full DSM criteria) for hypomania, mania or mixed state, or depression is more specific, but far less sensitive,

than criteria based on fewer symptoms or shorter durations than syndromal definitions require. Scales such as the YMRS or MADRS can be used with absolute cutoff scores (> 8 or > 15), but may be overly sensitive and inflate the switch; hence, they generate pseudo switch rates.

Most studies that specify a definition for affective switch use a single window, such as 8–12 weeks from the last change in the causal treatment. This criterion is reasonable in that it is easily established and corresponds to the time required for an antidepressant or antipsychotic to exert its psychotropic effects, but it ignores the high potential for false positives when applied to patients with cycle frequencies of three months or less, as well as false negatives if the presence of an agent with a liability to TEAS is enduring. The usefulness of attributing causality to specific treatments such as antidepressants has no doubt clinical utility; therefore, the task force recommends further consideration to include the specific treatment if the switch emerges within less than two weeks and to refer to it as an antidepressant-associated TEAS.

Summary of task force recommendations (see Table 7)

We recommend consideration of the term ‘TEAS’, as well as the development of operational definitions that consider causality (unlikely; possible; likely; definite), amplitude, duration, and window from intervention (last change in treatment). Clinically relevant methodology to validate TEAS to be considered includes its ability to predict relapse/recurrence over a period of 12 months.

**Functionality in bipolar disorder**

In recent years, it has become apparent that, in addition to the functional deficits that accompany both the depressive and manic phases of bipolar disorder, social, occupational, and cognitive functioning is often severely compromised even during periods of euthymia. The source of these functional deficits during euthymia has been difficult to determine and it is still not clear to what extent they reflect trait neuropathology, the vestigial effects of mood state changes, the effects of

Table 6. Predominant polarity

Rationale	• Predominant polarity of mood episodes of patients’ lifelong history may have clinical relevance for their therapeutic management
Task force recommendation	• Either predominantly depressed (PD) or predominantly manic (PM), defined by having at least two-thirds of lifetime episodes of one polarity or the other. The inclusion of a predominately mixed category is currently not recommended
Validating methodology	• Ability to predict relapse/recurrence over a 12-month period and the polarity of the relapse/recurrence

Table 7. (A) Treatment-emergent affective switch (TEAS)

Rationale	<ul style="list-style-type: none"> <li>• Causality should not be a categorical ‘yes’ or ‘no’, but likelihood based on specified criteria</li> <li>• The term ‘TEAS’ does not attribute causality, and is preferred to ‘antidepressant-induced switch’ or ‘antipsychotic-induced switch’</li> </ul>
Task force recommendation	<ul style="list-style-type: none"> <li>• Consider use of the term TEAS and operational definitions based on causality, amplitude, duration, and window from intervention</li> </ul>
Validating methodology	<ul style="list-style-type: none"> <li>• Ability to predict relapse/recurrence over a 12-month period</li> </ul>

(B) Definitions

Causality	Amplitude	Duration	Window from intervention (last change)
<b>Treatment-emergent manic switch</b>			
Definite	Full syndromic hypomanic, manic, or mixed episode	At least 2 consecutive days with daily occurrence of symptomatic periods lasting more than 50% of time each day	≤ 8 weeks, if ≤ 2 weeks consider specifying type of treatment (e.g., antidepressant-associated)
Likely	2 or more symptoms (e.g., irritability or euphoria, racing thoughts, grandiosity, decreased need for sleep) and YMRS score of > 12	At least 2 consecutive days with daily occurrence of symptomatic periods lasting more than 50% of time each day	≤ 12 weeks
Possible	Clear change in mood or energy and YMRS score of > 8	Symptomatic periods each day lasting at least 4 hours over a 2-day period	≤ 12 weeks
Unlikely	Fleeting symptoms, identifiable environmental or exogenous contribution		> 16 weeks
<b>Treatment-emergent depressive switch</b>			
Definite	Full syndromic depression	At least one consecutive week with daily occurrence of symptomatic periods lasting more than 50% of time each day	≤ 8 weeks, if ≤ 2 weeks consider specifying type of treatment (e.g., antipsychotic-induced)
Likely	2 or more depressive symptoms and MADRS score of > 12	At least one consecutive week with daily occurrence of symptomatic periods lasting more than 50% of time each day	≤ 12 weeks
Possible	Clear change in mood or energy and MADRS score of > 8	Symptomatic periods each day lasting at least 4 hours over a one-week period	≤ 12 weeks
Unlikely	Fleeting symptoms, identifiable environmental or exogenous contribution		> 16 weeks

YMRS = Young Mania Rating Scale; MADRS = Montgomery-Åsberg Depression Rating Scale.

untreated comorbid psychiatric conditions such as panic disorder, comorbid medical conditions such as obesity and cardiovascular disease, the side effects of ongoing medication, or some combination of the above (43, 78). Deficits in functioning have been identified even after the first episode (7, 40).

Following acute episodes of illness, residual symptoms or subsyndromal states often persist, sometimes throughout the interepisode intervals of the disorder (57). The nature of these symptoms, along with factors such as age of onset and pattern of illness (in particular, the frequency of episodes and presence of cycling and mixed states), clearly have an impact on functional outcome in bipolar disorder. Comorbidity and adherence to treatment no doubt also play a role; however, it is the social, occupational, and neurocognitive profile of patients with bipolar disorder that needs to be delineated to fully appreciate the impact the illness

has on functioning when patients are otherwise seemingly well (78).

Poor social functioning in bipolar disorder is common (79, 80); however, the factors contributing to this effect have been found to vary, with only depressive symptoms emerging as a consistent and reliable predictor of social adjustment. Historically, demographic factors have been weighted more heavily than clinical variables (81, 82), and more recent studies that have examined personality in conjunction with clinical symptoms have found that neuroticism alongside depression is predictive of overall quality of social functioning (83). Subsyndromal, interepisodic depressive states are also important (7, 57), as they impact upon the quality of social functioning. Interestingly, the link between functional impairment and illness in a relapsing and recurrent disorder such as bipolar disorder is most likely to be iterative. Functional

compromise often also precedes the onset of illness episodes, indicating perhaps that there is some cognitive vulnerability to relapse (84). Thus, it is important to note that although mood is a significant factor in determining social adjustment, it is not the overriding factor. A mix of demographic and clinical variables such as personality disorder (borderline/antisocial), male gender, isolation, and use of medications for sleep or physical illnesses also contribute significantly to social adjustment (84). A better understanding of these relationships and their respective contributions is important, as some factors, such as mood, are clearly amenable to treatment and therapeutic intervention, whereas the effects of others, such as gender, personality disorder, and social circumstances, are more difficult to modify from the clinical perspective.

In recent years, a more widely adopted framework for classifying the range of problems that patients with bipolar disorder experience with respect to functioning has been developed [the International Classification of Functioning, Disability and Health (ICF)]. This takes into consideration environmental and psychosocial perspectives, and partitions effects into three components: body structures and functions; activities and participation; and personal environmental factors. The breadth of the components encompassed by the ICF allows tailoring to specific disorders, and in recent years a number of researchers have attempted to develop an ICF core set for bipolar disorder (84). In practice, an assessment tool is needed to describe and classify functioning in patients with bipolar disorder that can be administered by clinicians and researchers alike. Such an instrument could become the basis of assessments that inform treatment and long-term management. Apart from its clinical utility across patient populations, a standardized tool would also allow more finely focused research within this domain. However, in order to have widespread value, any such tool needs to be easily administered with high reliability.

It is important to note that the neurocognitive deficits in bipolar disorder that contribute to and perhaps underpin impairment in psychosocial functioning impairment are different than those found in schizophrenia (85). Cognitive deficits in

bipolar disorder across its various mood states are well recognized and have been reliably documented in nearly 50 studies (78). Euthymic bipolar patients have been shown to have deficits across these same domains, encompassing cognitive flexibility, attention and language tasks, and in particular verbal memory (86–88). Interestingly, it is the extent and degree of cognitive compromise that appears to be related to the degree of functioning (86), with low-functioning patients much more likely to possess significant cognitive impairments. In this regard, verbal memory dysfunction has been shown to be a putative predictor of psychosocial outcome in euthymic bipolar disorder.

The assessment of functioning is therefore necessarily complex and requires a sophisticated instrument that is simple to administer. One such assessment that has been recently developed is the Functioning Assessment Short Test (FAST) (89). The FAST is a 24-item scale that spans six domains of functioning: autonomy, occupational functioning, cognitive functioning, financial issues, interpersonal relationships, and leisure time. It takes approximately five minutes to complete the scale, which focuses on the principal problems faced by bipolar patients. It can be administered during any of the mood states and is designed to detect change. It is also suitable for both clinical and research settings and can be used longitudinally to examine the effects of medication and other interventions. It does not, however, measure QOL, and it is clinician administered. These limitations can perhaps be overcome in instances where these parameters are important by using it alongside self-report and performance-based instruments that also capture QOL. A recently developed, novel way of describing course and outcome which includes functioning and has been utilized in other areas of medicine, known as Staging Models, has recently been proposed for use in bipolar disorders (90). Additional studies to validate this terminology are further needed.

#### Summary of task force recommendations (see Table 8)

Functional assessment in bipolar disorder cannot be captured accurately by conventional mood

Table 8. Functionality

Rationale	<ul style="list-style-type: none"> <li>• Although the nature of subsyndromal states, age of onset, pattern of illness, comorbidity, and adherence to treatment play a role in outcome, the neurocognitive profile influences the impact of illness on functioning in seemingly well patients</li> </ul>
Task force recommendation	<ul style="list-style-type: none"> <li>• Consider use of the Functioning Assessment Short Test (FAST) focused on principal problems of patients with bipolar disorder and spanning 6 domains of functioning: autonomy, occupational functioning, cognitive functioning, financial issues, interpersonal relationships, and leisure time</li> </ul>
Validating methodology	<ul style="list-style-type: none"> <li>• Neurocognitive processes</li> </ul>



rating scales. Recent research has shown that clinical functioning is related to neurocognitive processes; however, this aspect requires further research to determine the precise relationship. A recent instrument (FAST) has been developed to capture function in bipolar disorder, and others will emerge imminently. Widespread use and standardization of these tools is needed, perhaps along with concurrent use of other scales, in order to realize the full impact of psychosocial compromise in bipolar disorder patients.

### Conclusions

The content of this article is for the most part a consensus opinion based on the existing literature on definitions of terms frequently used in clinical trials and observational study. For the most part, the definitions have not been validated. In proposing consensus definitions of course and outcome in bipolar disorder, we have emphasized rating scale criteria in several sections of this article, as these quantifiable approaches can be compared across studies and provide important criteria for inclusion, exclusion, and operationalized outcome definitions for future clinical trials and observational studies. Most clinicians do not routinely apply such scales. We hope that clinicians will apply the principles embodied here in their assessment and clinical treatment of patients, even if most of their information is qualitative. Perhaps this approach can be most usefully applied by considering what may be most critical, i.e., what should not be dismissed when making diagnostic and treatment decisions, given the vexing symptomatic overlaps and sometimes arbitrary durational requirements in the current DSM-IV classification system. Further refinement of current classifications, namely those of the DSM-V and ICD-11, might include a dimensional assessment of symptoms as an adjunct to the categorical system, which could be used as a tool to assess subthreshold states (91). However, diagnostic schema (whether DSM, ICD, or another) cannot be expected to provide an operational tool to guide clinical management across time. This report represents a first step in the development of an empirically derived nomenclature on the course and outcome of bipolar disorder. The next step is to further develop the methodology to validate the task force recommendations, in both observational studies and clinical trials, in order to develop definitions that are clinically meaningful and helpful in the care of patients suffering from bipolar disorder.

### Acknowledgements

This work was conducted under the auspices of The International Society for Bipolar Disorders. Editorial assistance was provided by Caron Modeas.

### Affiliations

<sup>a</sup>Department of Psychiatry, University of Texas Health Science Center at San Antonio, San Antonio, TX, <sup>b</sup>McLean Hospital, Harvard Medical School, Belmont MA, <sup>c</sup>Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA, <sup>d</sup>Department of Psychiatry, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain, <sup>e</sup>Department of Psychiatry, Tufts University, Boston MA, USA, <sup>f</sup>Department of Psychiatry, Division of Mood Disorders, University of British Columbia, Vancouver, BC, Canada, <sup>g</sup>CADE Clinic, Department of Psychiatry; and Discipline of Psychological Medicine, University of Sydney, Sydney, Australia, <sup>h</sup>Department of Psychiatry, Case Western Reserve University School of Medicine/University Hospitals of Cleveland, Cleveland, OH, USA, <sup>i</sup>Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands, <sup>j</sup>Department of Psychiatry, Federal University, UFRGS, HCPA, INCT-Translational Medicine, Porto Alegre, Brazil, <sup>k</sup>Department of Psychiatry, Warneford Hospital, Oxford, UK, <sup>l</sup>Department of Psychiatry, Stanford University, Palo Alto, CA, <sup>m</sup>Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA, <sup>n</sup>Department of Psychiatry, University of Newcastle Institute of Neuroscience, Royal Victoria Infirmery, Newcastle upon Tyne, UK, <sup>o</sup>School of Psychiatry, University of New South Wales, and Black Dog Institute, Sydney, Australia, <sup>p</sup>Department of Neuropsychiatry, Kyushu University, Fukuoka, Japan, <sup>q</sup>Department of Clinical and Biomedical Sciences, Melbourne University, Melbourne, Australia

### Disclosures

MT is a former employee of Eli Lilly & Co. (2008) and his spouse is a current Eli Lilly & Co. employee and stockholder; and he has been a consultant to or received honoraria from AstraZeneca, Bristol-Myers Squibb, Eli Lilly & Co., GlaxoSmithKline, Johnson & Johnson and Wyeth. EF has received grant/research support from the Fine Foundation, the Pittsburgh Foundation, and Forest Research Institute; consultant fees from Eli Lilly & Co., Novartis, Lundbeck, Servier and Pfizer; and royalties from Guilford Press. CLB has received research/grant support from Abbott, GlaxoSmithKline, Janssen, Eli Lilly & Co., and NIMH; and has served as a consultant to Abbott, Bristol-Myers Squibb, Eli Lilly & Co., Pfizer, and Sanofi-aventis. FC has served on the speakers bureau for AstraZeneca, GlaxoSmithKline, Pfizer, Sanofi-aventis, Eli Lilly & Co., and Tecnofar; and has served on the advisors panel for Shire and AstraZeneca. SNG has received research grants from Pfizer; has served on the speakers bureaus of AstraZeneca and Pfizer; and has received honoraria from Bristol-Myers Squibb; neither he nor his family holds equity positions in pharmaceutical corporations. LNY has received research grants from or is on the speakers/advisory boards for AstraZeneca, Bristol-Myers Squibb, Canadian Institutes of Health Research, Canadian Network for Mood and Anxiety Treatments, Eli Lilly & Co., GlaxoSmithKline, Janssen, Michael Smith Foundation for Health Research, Novartis,

Pfizer, Ranbaxy, Servier, and the Stanley Foundation. GSM has received funding for investigator-initiated research from Eli Lilly & Co., AstraZeneca, and Wyeth; has served on the advisory board of Eli Lilly & Co., AstraZeneca, Wyeth, and Pfizer; and has received honoraria for presentations from Organon, Eli Lilly & Co., AstraZeneca, Wyeth, Pfizer, Sanofi-aventis, and Janssen-Cilag; however, no part of this manuscript was discussed with any pharmaceutical company by GSM. JRC has received research/grant support from the Department of Defense, the Health Resources Services Administration, NIMH, the Cleveland Foundation, NARSAD, the Stanley Medical Research Institute, Abbott, AstraZeneca, Cephalon, GlaxoSmithKline, Janssen, Eli Lilly & Co., and Lundbeck; has served on the advisory boards of Abbott, AstraZeneca, Bristol-Myers Squibb, Dainippon Sumitomo Pharma, Forest Labs, France Foundation, GlaxoSmithKline, Janssen, NeuroSearch, Schering-Plough, OrthoMcNeil, Repligen, Servier, Solvay, Supernus Pharmaceuticals, Takeda, Wyeth; and has performed CME with AstraZeneca, Bristol-Myers Squibb, France Foundation, GlaxoSmithKline, Janssen/Johnson & Johnson, Schering-Plough, Sanofi-aventis, Solvay, and Wyeth. WAN has received research/grant support from the Netherlands Organisation for Health Research and Development, European Union, the Stanley Medical Research Institute, AstraZeneca, Eli Lilly & Co., GlaxoSmithKline, and Wyeth; has received honoraria/speakers' fees from AstraZeneca, Eli Lilly & Co., Pfizer, Servier, and Wyeth; and has served on the advisory boards of AstraZeneca, Cyberonics, Pfizer, and Servier. EV has received grants from and/or served as consultant, advisor, or speaker for Almirall, AstraZeneca, Bristol-Myers Squibb, Eli Lilly & Co., Forest, GlaxoSmithKline, Janssen-Cilag, Jazz, Lundbeck, Merck, Novartis, Organon, Otsuka, Pfizer, Sanofi-aventis, Servier, Shering-Plough, the Spanish Ministry of Science and Innovation (CIBERSAM), the Seventh European Framework Programme (ENBREC), the Stanley Medical Research Institute, United Biosource Corporation, and Wyeth. FK has received grant support from CNPq, CAPES, INCT, NARSAD, and the Stanley Medical Research Institute; and has served as an advisor/speaker for Abbott, AstraZeneca, Eli Lilly & Co., Janssen-Cilag, and Servier. GMG has held grants from Sanofi-aventis and Servier; and has accepted honoraria for speaking or advice from AstraZeneca, Bristol-Myers Squibb, Eisai, Eli Lilly & Co., Lundbeck, Plvital, Sanofi-aventis, Servier, and Wyeth. TS has received research funding/grants from Abbott, AstraZeneca, GlaxoSmithKline, JDS Pharmaceuticals, NIMH, Novartis, Pfizer, and the Stanley Medical Research Institute; and has a consulting agreement/served on the advisory board for Orexigen Therapeutics, Inc.; and has received royalties from Compact Clinicals. GSS has received research support and/or served on the speakers bureau or advisory boards for Abbott, AstraZeneca, Bristol-Myers Squibb, Cephalon, Dainippon Sumitomo Pharma, Eli Lilly & Co, GlaxoSmithKline, Janssen, Memory Pharmaceuticals, Merck, NIMH, Novartis, Otsuka, Pfizer, Repligen, Sanofi-aventis, Schering-Plough, Sepacor, Shire, Solvay and Wyeth; and has equity in Concordant Rater Systems. KNRC has received funding for an investigator-initiated study from Janssen-Ortho; grant support from the Stanley Medical Research Institute, NARSAD, and NIDA; and has received honoraria for speaking engagements and/or consulting from Eli Lilly & Co., AstraZeneca, and Janssen. HG has received research support from AstraZeneca, UCB Belgium, the Stanley Foundation, and the NHS National Institute for Health Research/Medical Research Council UK; and has served as a consultant/advisory board member or speaker for AstraZeneca, Bristol-Myers Squibb, Eli Lilly & Co., GlaxoSmithKline, Janssen-Cilag, Pfizer, Sanofi-aventis, and

UBC. PBM has received remuneration for advisory board membership from Eli Lilly & Co. and AstraZeneca; and consultant fees or lecture honoraria from AstraZeneca, Eli Lilly & Co., Janssen-Cilag, and Lundbeck; however, he is currently not a member of any pharmaceutical company advisory board. SK has received research/grant support from Eli Lilly & Co., GlaxoSmithKline, Pfizer, Asahi-kasei, Janssen, Tsumura, Ajinomoto, Yoshitomi, Meiji, Kyowa-Hakko, Sumitomo, Organon, and Otsuka; has served as consultant for Eli Lilly & Co., GlaxoSmithKline, Pfizer, Mitsubishi, Ono, Astellas, Asahi-kasei, Shionogi, and Otsuka; and has received honoraria from Eli Lilly & Co., GlaxoSmithKline, Pfizer, Asahi-kasei, Janssen, Tsumura, Ajinomoto, Yoshitomi, Meiji, Kyowa-Hakko, Sumitomo, Organon, Otsuka, and Astellas. MB has received grant/research support from the Stanley Medical Research Foundation, MBF, NHMRC, Beyond Blue, the Geelong Medical Research Foundation, Bristol-Myers Squibb, Eli Lilly & Co., GlaxoSmithKline, Organon, Novartis, Mayne Pharma, and Servier; has served as a speaker for AstraZeneca, Bristol-Myers Squibb, Eli Lilly & Co., GlaxoSmithKline, Janssen Cilag, Lundbeck, Pfizer, Sanofi Synthelabo, Servier, Solvay, and Wyeth; and has served as a consultant for AstraZeneca, Bristol-Myers Squibb, Eli Lilly & Co., GlaxoSmithKline, Janssen Cilag, Lundbeck, and Servier.

## References

1. Rush AJ, Kraemer HC, Sackeim HA et al. Report by the ACNP task force on response and remission in major depressive disorder. *Neuropsychopharmacology* 2006; 31: 1841–1853.
2. Tohen M, Bowden C, Calabrese J et al. The International Society of Bipolar Disorders task force on the nomenclature of course and outcome in bipolar disorders. *Bipolar Disord* 2008; 10 (Suppl. 1): 1.
3. Andreasen NC, Carpenter WT Jr, Kane JM, Lasser RA, Marder SR, Weinberger DR. Remission in schizophrenia: proposed criteria and rationale for consensus. *Am J Psychiatry* 2005; 162: 441–449.
4. Leucht S, Davis JM, Engel RR, Kane JM, Wagenpfeil S. Defining 'response' in antipsychotic drug trials: recommendations for the use of scale-derived cutoffs. *Neuropsychopharmacology* 2007; 32: 1903–1910.
5. Frank E, Prien RF, Jarrett JB et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: response, remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry* 1991; 48: 851–855.
6. American Psychiatric Association. *Diagnostic and Statistical Manual for Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*. Washington, DC: American Psychiatric Press, 2000.
7. Tohen M, Zarate CA, Hennen J et al. The McLean-Harvard first-episode mania study: prediction of recovery and first recurrence. *Am J Psychiatry* 2003; 160: 2099–2107.
8. Hirschfeld RM, Calabrese J, Frye MA et al. Defining the clinical course of bipolar disorder response, remission, relapse, recurrence, and roughening. *Psychopharmacol Bull* 2007; 40: 7–14.
9. Martínez-Arán A, Vieta E, Chengappa KNR, Gershon S, Mullen J, Paulsson B. Reporting outcomes in clinical trials for bipolar disorder: a commentary and suggestions for change. *Bipolar Disord* 2008; 10: 566–579.
10. Yatham L, Vieta E, Berk M, Tohen M. Moving the field forward: the impact of research driven educational initiatives on bipolar patients and providers. Presented at the

- 3rd Biennial Conference of the International Society of Affective Disorders, Lisbon, Portugal, 3–6, March 2006.
11. Fava M, Evins AE, Dorer DJ, Schoenfeld DA. The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychother Psychosom* 2003; 72: 115–127.
  12. Rush AJ, Thase ME, Dube S. Research issues in the study of difficult-to-treat depression. *Biol Psychiatry* 2003; 53: 743–753.
  13. Vieta E, Jaeger J. Functional outcome and disability in bipolar disorders: ongoing research and future directions. *Bipolar Disord* 2007; 9: 1–2.
  14. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; 23: 56–62.
  15. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; 134: 382–389.
  16. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS). Psychometric properties. *Psychol Med* 1996; 26: 477–486.
  17. Berk M, Malhi GS, Cahill C et al. The Bipolar Depression Rating Scale (BDRS): its development, validation and utility. *Bipolar Disord* 2007; 9: 571–579.
  18. Young RC, Biggs JT, Ziegler VE, Merger DA. A rating scale for mania. Reliability, validity, sensitivity. *Br J Psychiatry* 1979; 133: 429–433.
  19. Guy W. Clinical global impressions. In: Guy W, ECDEU Assessment Manual for Psychopharmacology, Revised (*DHEW Publ No ADM 76-338*). Rockville, MD: National Institute of Mental Health, 1976: 218–222.
  20. Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychol Rep* 1962; 10: 790–812.
  21. Kendler KS. Major depression and generalized anxiety disorder. Same genes, (partly) different environments—revisited. *Br J Psychiatry* 1996; 168 (Suppl. 30): 68–75.
  22. Endicott J, Spitzer RL. A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry* 1978; 35: 837–844.
  23. Lecrubier Y, Sheehan D, Weiller E et al. The MINI International Neuropsychiatric Interview (M.I.N.I.). A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur Psychiatry* 1997; 12: 224–231.
  24. Kroenke K, Spitzer RL, Williams JB. The PHQ-9. Validity of a brief depression severity measure. *J Gen Intern Med* 2001; 16: 606–613.
  25. Rush AJ, Trivedi MH, Ibrahim HM et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS) Clinician Rating (QIDS-C) and Self-Report (QIDS-SR). A psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 2003; 54: 573–583.
  26. Trivedi MH, Rush AJ, Ibrahim HM et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med* 2004; 34: 73–82.
  27. Beck AT, Ward CH, Mendelson M, Mock JE, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961; 4: 561–571.
  28. Beck AT, Steer RA, Brown GK. Beck depression inventory (2nd ed. Manual). San Antonio, TX: The Psychological Corporation, 1996.
  29. Nierenberg AA, Keefe BR, Leslie VC et al. Residual symptoms in depressed patients who respond acutely to fluoxetine. *J Clin Psychiatry* 1999; 60: 221–225.
  30. Carmody T, Rush AJ, Bernstein I et al. The Montgomery Åsberg and the Hamilton Ratings of depression: a comparison of measures. *Eur Neuropsychopharmacol* 2006; 16: 601–611.
  31. Zimmerman M, Chelminski I, Posternak M. A review of studies of the Montgomery-Åsberg Depression Rating Scale in controls: implications for the definition of remission in treatment studies of depression. *Int Clin Psychopharmacol* 2004; 19: 1–7.
  32. Zimmerman M, Posternak MA, Chelminski I. Defining remission on the Montgomery-Åsberg depression rating scale. *J Clin Psychiatry* 2004; 65: 163–168.
  33. Zimmerman M, Posternak MA, Chelminski I. Implications of using different cut-offs on symptom severity scales to define remission from depression. *Int Clin Psychopharmacol* 2004; 19: 215–220.
  34. 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: 153–158.
  35. Post R, Altshuler LL, Leverich GS et al. Mood switch in bipolar depression: comparison of adjunctive venlafaxine, bupropion and sertraline. *Br J Psychiatry* 2006; 189: 124–131.
  36. Frye MA, Grunze H, Suppes T et al. A placebo-controlled evaluation of adjunctive modafinil in the treatment of bipolar depression. *Am J Psychiatry* 2007; 164: 1242–1249.
  37. Thase ME, Carpenter L, Kupfer DJ, Frank E. Clinical significance of reversed vegetative subtypes of recurrent major depression. *Psychopharmacol Bull* 1991; 27: 17–22.
  38. Chengappa KNR, Hennen J, Baldessarini RJ et al. Recovery and functional outcomes following olanzapine treatment for bipolar I mania. *Bipolar Disord* 2005; 7: 68–76.
  39. Keller MB, Lavori PW, Coryell W, Endicott J, Mueller TI. Bipolar I: a five year prospective follow-up. *J Nerv Ment Dis* 1993; 181: 238–245.
  40. Tohen M, Hennen J, Zarate C et al. The McLean/Harvard First Episode Project: two-year syndromal and functional recovery in 219 cases of major affective disorders with psychotic features. *Am J Psychiatry* 2000; 157: 220–228.
  41. Perlis RH, Ostacher MJ, Patel JK et al. Predictors of recurrence in bipolar disorder: primary outcomes from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *Am J Psychiatry* 2006; 163: 217–224.
  42. Riso LP, Thase ME, Howland RH, Friedman ES, Simons AD, Tu XM. A prospective test of criteria for response remission relapse recovery and recurrence in patients treated with cognitive behavior therapy. *J Affect Disord* 1997; 43: 131–142.
  43. Tohen M, Waternaux CM, Tsuang MT. Outcome in mania. A four-year prospective follow-up of 75 patients utilizing survival analysis. *Arch Gen Psychiatry* 1990; 47: 1106–1111.
  44. Ghaemi SN, Pardo TB, Hsu DJ. Strategies for preventing the recurrence of bipolar disorder. *J Clin Psychiatry* 2004; 65 (Suppl. 10): 16–23.
  45. Goodwin FK, Jamison KR. *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression*, 2nd ed. New York: Oxford University Press, 2007.
  46. Hartong EG, Moleman P, Goodguin CA, Broekman TG, Nolen WA, LitCar Group. Prophylactic efficacy of lithium versus carbamazepine in treatment-naïve bipolar patients. *J Clin Psychiatry* 2003; 64: 144–151.

47. Tohen M, Greil W, Calabrese JR et al. Olanzapine versus lithium in the maintenance treatment of bipolar disorder: a 12-month, randomized, double-blind, controlled clinical trial. *Am J Psychiatry* 2005; 162: 1281–1290.
48. Greil W, Ludwig-Mayerhofer W, Erazo N et al. Lithium versus carbamazepine in the maintenance treatment of bipolar disorders—a randomised study. *J Affect Disord* 1997; 43: 151–161.
49. Bowden CL, Calabrese JR, McElroy SL et al. A randomized, placebo-controlled 12-month trial of divalproex and lithium in treatment of outpatients with bipolar I disorder. Divalproex Maintenance Study Group. *Arch Gen Psychiatry* 2000; 57: 481–489.
50. Tohen M, Calabrese JR, Sachs GS et al. Randomized, placebo-controlled trial of olanzapine as maintenance therapy in patients with bipolar I disorder responding to acute treatment with olanzapine. *Am J Psychiatry* 2006; 163: 247–256.
51. Calabrese JR, Shelton MD, Rapport DJ et al. A 20-month, double-blind, maintenance trial of lithium versus divalproex in rapid-cycling bipolar disorder. *Am J Psychiatry* 2005; 162: 2152–2161.
52. Bowden CL, Calabrese JR, Sachs G et al. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I disorder. *Arch Gen Psychiatry* 2003; 60: 392–400.
53. Calabrese J, Bowden C, Sachs G et al. A placebo-controlled, 18-month trial of lamotrigine and lithium maintenance treatment in recently depressed patients with bipolar I disorder. *J Clin Psychiatry* 2003; 64: 1013–1024.
54. Judd LL, Akiskal HS, Schettler PJ et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch Gen Psychiatry* 2002; 59: 530–537.
55. Altshuler LL, Gitlin MJ, Mintz J, Leight KL, Frye MA. Subsyndromal depression is associated with functional impairment in patients with bipolar disorder. *J Clin Psychiatry* 2002; 63: 807–811.
56. MacQueen GM, Marriott M, Begin H, Robb J, Joffe RT, Young LT. Subsyndromal symptoms assessed in longitudinal prospective follow-up of a cohort of patients with bipolar disorder. *Bipolar Disord* 2003; 5: 349–355.
57. Tohen M, Bowden CL, Calabrese JR et al. Influence of sub-syndromal symptoms after remission from manic or mixed episodes. *Br J Psychiatry* 2006; 189: 515–519.
58. Torrent C, Martínez-Arán A, Daban C et al. Cognitive impairment in bipolar II disorder. *Br J Psychiatry* 2006; 189: 254–259.
59. Yatham LN, Lecrubier Y, Fieve RR, Davis KH, Harris SD, Krishnan AA. Quality of life in patients with bipolar I depression: data from 920 patients. *Bipolar Disord* 2004; 6: 379–385.
60. Marangell LB. The importance of subsyndromal symptoms in bipolar disorder. *J Clin Psychiatry* 2004; 65 (Suppl. 10): 24–27.
61. Bagby R, Ryder AG, Schuller DR, Marshall MB. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *Am J Psychiatry* 2004; 161: 2163–2177.
62. Cassidy F, Pieper CF, Carroll BJ. Subtypes of mania determined by grade of membership analysis. *Neuropsychopharmacology* 2001; 25: 373–383.
63. Dennehy EB, Suppes T, Crismon ML, Toprac M, Carmody TJ, Rush AJ. Development of The Brief Bipolar Disorder Symptom Scale for patients with bipolar disorder. *Psychiatry Res* 2004; 127: 137–145.
64. Swann CA, Janicak LP, Calabrese RJ et al. Structure of mania: depressive, irritable, and psychotic cluster with different retrospectively-assessed course patterns of illness in randomized clinical trial participants. *J Affect Dis* 2001; 67: 123–132.
65. Kessler R, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005; 62: 593–602.
66. Bowden C, Singh V, Thompson P et al. Development of the bipolar inventory of symptoms scale. *Acta Psychiatr Scand* 2007; 116: 189–194.
67. Fagiolini A, Dell’Osso L, Pini S et al. Validity and reliability of a new instrument for assessing mood symptomatology: the Structured Clinical Interview for Mood Spectrum (SCI-MOODS). *Int J Methods Psychiatr Res* 1999; 8: 71–82.
68. 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.
69. Miniati M, Rucci P, Frank E et al. Sensitivity to change and predictive validity of the MOODS-SR questionnaire, last-month version. *Psychother Psychosom* 2009; 78: 116–124.
70. Judd LL, Schettler PJ, Akiskal HS et al. Residual symptom recovery from major affective episodes in bipolar disorders and rapid episode relapse/recurrence. *Arch Gen Psychiatry* 2008; 65: 386–394.
71. Colom F, Vieta E, Daban C, Pacchiarotti I, Sanchez-Moreno J. Clinical and therapeutic implications of predominant polarity in bipolar disorder. *J Affect Disord* 2006; 93: 13–17.
72. Rosa AR, Andreazza AC, Kunz M et al. Predominant polarity in bipolar disorder: diagnostic implications. *J Affect Disord* 2008; 107: 45–51.
73. Vieta E, Berk M, Wang W, Colom F, Tohen M, Baldessarini RJ. Predominant previous polarity as an outcome predictor of outcome in a controlled treatment trial for depression in bipolar I disorder patients. *J Affect Disord*, [E pub ahead of print, 2009 March 24], doi: 10.1016/j.jad.2009.02.028.
74. Osher Y, Yaroslavsky Y, el-Rom R, Belmaker RH. Predominant polarity of bipolar patients in Israel. *World J Biol Psychiatry* 2000; 1: 187–189.
75. Daban C, Colom F, Sanchez-Moreno J, Garcia-Amador M, Vieta E. Clinical correlates of first-episode polarity in bipolar disorder. *Compr Psychiatry* 2006; 47: 433–437.
76. Judd LL, Akiskal HS, Schettler PJ et al. A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar II disorder. *Arch Gen Psychiatry* 2003; 60: 261–269.
77. Perlis RH, Delbello MP, Miyahara S et al. Revisiting depressive-prone bipolar disorder: polarity of initial mood episode and disease course among bipolar I systematic treatment enhancement program for bipolar disorder participants. *Biol Psychiatry* 2005; 58: 549–553.
78. Malhi GS, Ivanovski B, Szekeres V, Olley A. Bipolar disorder: it’s all in your mind? The neuropsychological profile of a biological disorder. *Can J Psychiatry* 2004; 49: 813–819.
79. Olley A, Malhi GS, Mitchell P, Batchelor J, Lagopoulos J, Austin M-P. When euthymia is just not good enough: the neuropsychology of bipolar disorder. *J Nerv Ment Dis* 2005; 193: 323–330.

## Nomenclature, bipolar disorder, course and outcome

80. MacQueen GM, Young LT, Joffe RT. A review of psychosocial outcome in patients with bipolar disorder. *Acta Psychiatr Scand* 2001; 103: 163–170.
81. Pope M, Dudley R, Scott J. Determinants of social functioning in bipolar disorder. *Bipolar Disord* 2007; 9: 38–44.
82. Goetz I, Tohen M, Reed C, Lorenzo M, Vieta E, EMBLEM Advisory Board. Functional impairment in patients with mania: baseline results of the EMBLEM study. *Bipolar Disord* 2007; 9: 45–52.
83. Morriss R, Scott J, Paykel E, Bentall R, Hayhurst H, Johnson T. Social adjustment based on reported behaviour in bipolar affective disorder. *Bipolar Disord* 2007; 9: 53–62.
84. Vieta E, Cieza A, Stucki G et al. Developing core sets for persons with bipolar disorder based on the International Classification of Functioning, Disability and Health. *Bipolar Disord* 2007; 9: 16–24.
85. Czobor P, Jaeger J, Berns SM, Gonzalez C, Loftus S. Neuropsychological symptom dimensions in bipolar disorder and schizophrenia. *Bipolar Disord* 2007; 9: 71–92.
86. Martínez-Arán A, Vieta E, Colom F et al. Cognitive impairment in euthymic bipolar patients: implications for clinical and functional outcome. *Bipolar Disord* 2004; 6: 224–232.
87. Malhi GS, Ivanovski B, Hadzi-Pavlovic D, Mitchell P, Vieta E, Sachdev P. Neuropsychological deficits and functional impairment in bipolar depression, hypomania and euthymia. *Bipolar Disord* 2007; 9: 114–125.
88. Martínez-Arán A, Vieta E, Torrent C et al. Functional outcome in bipolar disorder: the role of clinical and cognitive factors. *Bipolar Disord* 2007; 9: 103–113.
89. Rosa AR, Sanchez-Moreno J, Martínez-Arán A et al. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. *Clin Pract Epidemiol Ment Health* 2007; 7: 3–5.
90. Berk M, Hallam KT, McGorry PD. The potential utility of a staging model as a course specifier: a bipolar disorder perspective. *J Affect Disord* 2007; 100: 279–281.
91. Vieta E, Phillips ML. Deconstructing bipolar disorder: a critical review of its diagnostic validity and a proposal for DSM-V and ICD-11. *Schizophr Bull* 2007; 33: 886–892.