

Cumulative morbidity and prognostic staging of illness in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD)

Australian & New Zealand Journal of Psychiatry
00(0) 1–10
DOI: 10.1177/0004867412460593

© The Royal Australian and
New Zealand College of Psychiatrists 2012
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
anp.sagepub.com



Pedro Vieira Magalhães^{1,2}, Seetal Dodd^{2,3}, Andrew A Nierenberg⁴
and Michael Berk^{2,3,5,6}

Abstract

Objective: Staging models may provide heuristic utility for intervention selection in psychiatry. Although a few proposals have been put forth, there is a need for empirical validation if they are to be adopted. Using data from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD), we tested a previously elaborated hypothesis on the utility of using the number of previous episodes as a relevant prognostic variable for staging in bipolar disorder.

Methods: This report utilizes data from the multisite, prospective, open-label study ‘Standard Care Pathways’ and the subset of patients with acute depressive episodes who participated in the randomized trial of adjunctive antidepressant treatment. Outpatients meeting DSM-IV diagnostic criteria for bipolar disorder ($n = 3345$) were included. For the randomized pathway, patients met criteria for an acute depressive episode ($n = 376$). The number of previous episodes was categorized as less than 5, 5–10 and more than 10. We used disability at baseline, number of days well in the first year and longitudinal scores of depressive and manic symptoms, quality of life and functioning as validators of models constructed a priori.

Results: Patients with multiple previous episodes had consistently poorer cross-sectional and prospective outcomes. Functioning and quality of life were worse, disability more common, and symptoms more chronic and severe. There was no significant effect for staging with regard to antidepressant response in the randomized trial.

Conclusions: These findings confirm that bipolar disorder can be staged with prognostic validity. Stages can be used to stratify subjects in clinical trials and develop specific treatments.

Keywords

Bipolar disorder, cumulative morbidity, longitudinal, prognosis, staging

Introduction

In the past decade, there has been a new surge of interest in prognostic staging models in psychiatry, partly as a rationale for early intervention strategies (Francey et al., 2010; McGorry et al., 2007; Macneil et al., 2012; Raballo and Laroi, 2009). The fundamental idea behind clinical staging is that, akin to other medical disorders, it offers a link with pathological progression mechanisms. This places individuals in a continuum of the course of illness (McGorry, 2007). In this fashion, staging may be relevant to syndromes that tend to progress. This appears to be the case in bipolar disorder (Berk, 2009; Berk et al., 2010a, 2011b), where worse chances of recovery are apparent with cumulative burden (Solomon et al., 2010). As such,

¹National Institute for Translational Medicine, Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

²University of Melbourne, Department of Psychiatry, Parkville, Australia

³Deakin University, School of Medicine, Barwon Health, Geelong, Australia

⁴Massachusetts General Hospital, Harvard Medical School, Boston, USA

⁵Orygen Youth Health Research Centre, Parkville, Australia

⁶Florey Institute of Neuroscience and Mental Health, Parkville, Australia

Corresponding author:

Michael Berk, Department of Clinical and Biomedical Sciences: Barwon Health, University of Melbourne, P.O. Box 281, Geelong 3220, VIC, Australia.

Email: mikebe@barwonhealth.org.au

successful models refine response prediction and have the potential to improve treatment selection (Fava and Kellner, 1993; Vieta et al., 2011). As such, staging offers additional information on longitudinal course that complements the cross-sectional assessment of illness severity. As made explicit by McGorry and colleagues (McGorry et al., 2006), these models assume that early stages of illness have both better prognoses and a more benign and effective treatment.

Clinical stages can be defined in more than one manner. It has been argued that severity, course, and persistence of symptoms and their social impact, in addition to biological changes, should be part of the definition in psychiatry (McGorry et al., 2010). Building on that, more than one staging model has been put forward for bipolar disorder (Berk et al., 2007a, 2007b; Kapczinski et al., 2009; McGorry et al., 2006; Vieta et al., 2011). Even if not completely compatible, they all place multiple previous relapses at a late stage of the disorder. The number of previous episodes is not only an intuitive and pragmatic measure of cumulative illness burden, but is also germane to the current understanding of the recurrent and progressive nature of bipolar illness (Berk et al., 2011b; Kapczinski et al., 2008; Post, 2010). The little pathophysiological data available is consistent with this notion of a late stage with multiple previous relapses (Kapczinski et al., 2009; Kauer-Sant'Anna et al., 2009).

However interesting these frameworks are, they need to be assessed in longitudinal research. Verification of the prognostic value of stage of illness in independent populations is needed to validate a staging model (Biewenga et al., 2009; Harrell et al., 1996). The Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD), the largest treatment study of this illness conducted thus far (Sachs et al., 2003), provides a very large sample size with longitudinal evaluations, allowing for the testing of specific propositions.

The overall objective of this report was to confirm the hypothesis that patients with multiple previous relapses have an illness course that differs from individuals at an earlier disease stage. More explicitly, two assumptions are tested. Firstly, we used the STEP-BD Standard Care Pathways to test whether patients with multiple episodes had a worse symptomatic and functional prognosis. Secondly, we used data on the subset of patients that underwent the randomized placebo controlled trial (Sachs et al., 2007) to test the hypothesis that treatment with antidepressants would be more beneficial in the early stages of illness. If confirmed, this would strengthen the argument for treatments tailored to suit not only cross-sectional psychopathology, but also longitudinal course (McGorry et al., 2006).

Methods

Characterizations and thorough methodological descriptions of STEP-BD pathways have been extensively published

(Goldberg et al., 2009; Perlis et al., 2006, 2009, 2010; Sachs et al., 2003, 2007;). The multisite, prospective, open-label study was termed the 'Standard Care Pathways'. In this, participation was offered to those seeking outpatient treatment for bipolar disorder across sites and meeting DSM-IV diagnostic criteria for bipolar I disorder, bipolar II disorder, bipolar disorder not otherwise specified (NOS), cyclothymia or schizoaffective disorder, bipolar type. A subset of patients with acute depressive episodes during their participation was offered double-blind adjunctive antidepressant treatment. The study was approved by the human research committees of all treatment centers, and oral and written consent was obtained from participants.

Participants in open-label treatment could receive any intervention, as clinically indicated. Independent assessments were performed quarterly during the first year, which is the timeframe of this report. While the STEP-BD enrolled 4107 participants who had baseline data, the outcomes in this report include different sample sizes. Baseline, as well as data on number of days well in the first year, was available for 3345 individuals; for the longitudinal analyses, 2851 individuals had at least one post-baseline observation and were included in the analyses (1952 for the quality of life analysis). Retention over the study period varied, and sample sizes at the 12-month endpoint varied from 1516 participants for the MADRS and to 1015 for quality of life ratings.

The diagnosis of bipolar disorder was confirmed by a clinical rater with the Mini-International Neuropsychiatric Interview (MINI) (Sheehan et al., 1998), which was also used to obtain comorbid psychiatric diagnosis. Additional illness details were obtained with the Affective Disorders Evaluation (ADE) (Sachs et al., 2002). Information regarding number of previous episodes, age at onset, rapid cycling, history of psychotic episodes and current medication were obtained from this form.

For this report, number of previous episodes was categorized as less than 5, 5–10 and more than 10 previous episodes consistent with previous prospective studies on staging bipolar disorder (Berk et al., 2011a). In brief, this pooled analysis of olanzapine trials showed that people with five or fewer previous episodes had a better response to the treatment of acute depression ($n = 1.243$) and mania ($n = 1.631$); conversely, those with more than 10 episodes had a higher likelihood or relapse into mania or depression ($n = 1.432$). This study provided the empirical rationale for the categories employed here. As a secondary check, we performed the same analyses with variables representing more than 10 depressive and more than 10 manic episodes, with similar results (data not shown). We also created a childhood age at onset group, as it has recently been shown to be associated with worse outcomes in the STEP-BD (Perlis et al., 2009). History of clinical illness was obtained both from the ADE and Axis-III information that was coded from the dataset (Magalhaes et al., 2012a). Of interest in this category were chronic conditions that are prevalent and

have been associated with worse outcomes in bipolar disorder (Alonso et al., 2011; Kilbourne et al., 2008; Kupfer, 2005; Magalhaes et al., 2012b). These included cardiovascular conditions, diabetes, thyroid disorders, previous head injuries, migraine, epilepsy, multiple sclerosis, peptic ulcer and cancer. Subjects with any of these conditions were coded as having a medical comorbidity.

Interviewers assessed mood using the Montgomery-Asberg Rating Scale (Montgomery and Asberg, 1979) (MADRS) and the Young Mania Rating Scale (Young et al., 1978) (YMRS). Functioning and quality of life were assessed with the Longitudinal Interval Follow-up

Evaluation – Range of Impairment Functioning Tool (Keller et al., 1987) (LIFE-RIFT) and the Quality of life Enjoyment and Satisfaction Questionnaire (Endicott et al., 1993) (Q-LES-Q). We also analyzed a variable called ‘number of days well’, the number of days during the year of longitudinal study that patients were assigned a recovering or recovered clinical status. This was calculated with an algorithm that included clinical status during baseline and follow-up, YMRS and MADRS scores, and serious adverse events (Otto et al., 2006).

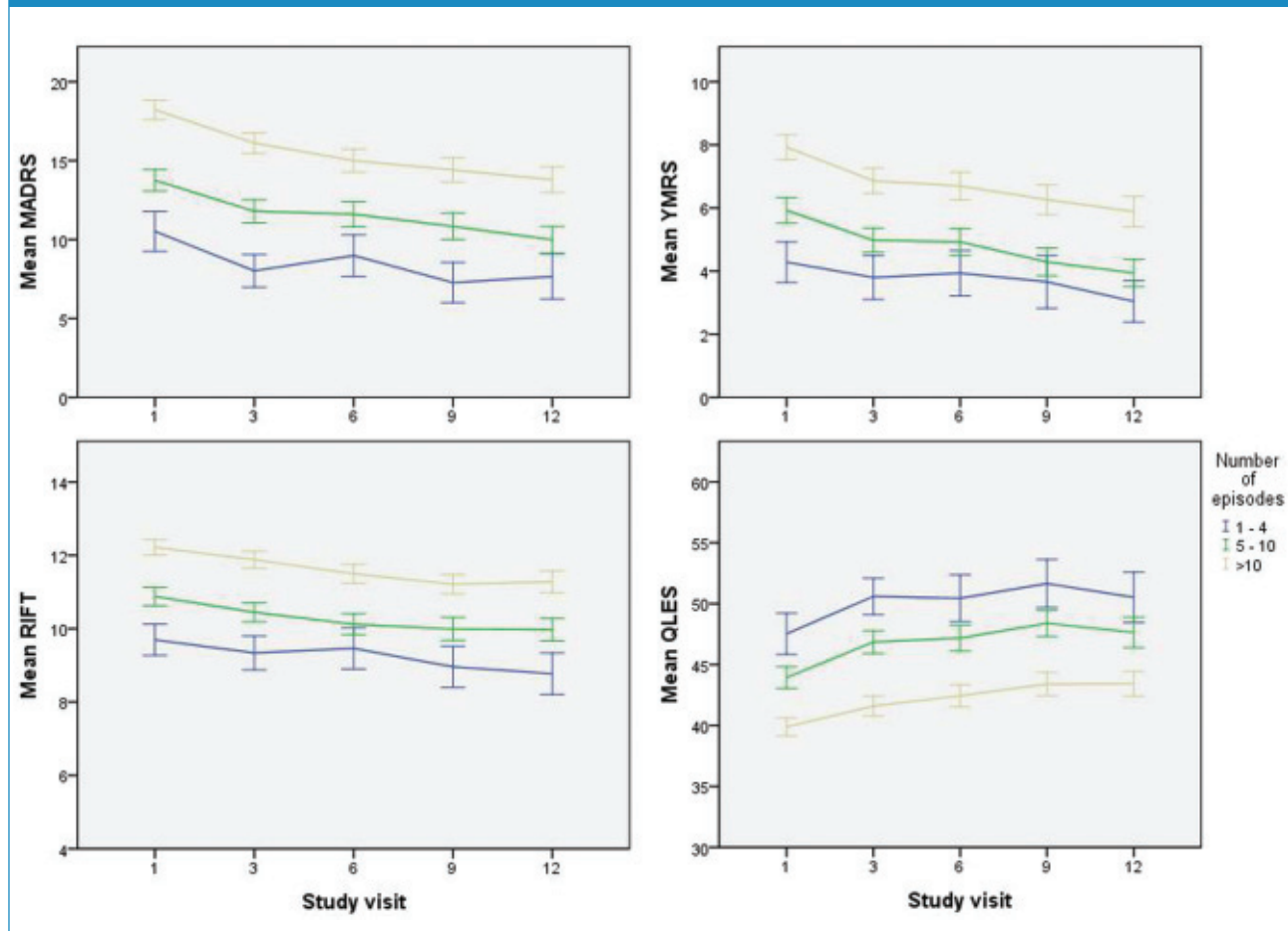
Patients in STEP-BD could be included in the randomized pathway if they fulfilled DSM-IV criteria for major

Table 1. Baseline characteristics of patients in the Standard Care Pathways according to number of previous episodes

| Characteristic | Total sample | Fewer than 5 episodes (<i>n</i> = 344) | Between 5 and 10 episodes (<i>n</i> = 1275) | More than 10 episodes (<i>n</i> = 1726) |
|--|--------------|---|--|--|
| Age* | 37.7± 12.9 | 36.9± 14.1 | 38.7± 13.3 | 40.9± 12.3 |
| Female sex* | 57% | 50% | 58% | 58% |
| Married or living with a partner | 34% | 33% | 34% | 34% |
| Low income (<\$US20,000 a year)* | 59% | 58% | 57% | 61% |
| On disability* | 15% | 8% | 14% | 20% |
| At least some college education* | 82% | 85% | 83% | 80% |
| Bipolar I disorder | 66% | 70% | 68% | 65% |
| Depressed at baseline* | 37% | 26% | 30% | 44% |
| Manic at baseline* | 5% | 3% | 4% | 6% |
| Rapid cycling* | 36% | 7% | 20% | 54% |
| Smoking at baseline | 33% | 23% | 30% | 37% |
| Lifetime psychosis* | 39% | 49% | 42% | 36% |
| Any anxiety disorder at baseline | 38% | 21% | 29% | 48% |
| Any substance use disorder at baseline | 17% | 12% | 15% | 19% |
| Childhood onset* | 29% | 12% | 19% | 39% |
| Any medical comorbidity* | 58% | 47% | 53% | 65% |
| Baseline meds | | | | |
| Lithium* | 31% | 35% | 34% | 28% |
| Other mood stabilizers | 56% | 54% | 57% | 57% |
| Atypical antipsychotics | 30% | 35% | 30% | 29% |
| Typical antipsychotics | 2% | 1% | 3% | 2% |
| Benzodiazepines | 24% | 16% | 20% | 28% |
| Antidepressants | 45% | 34% | 42% | 50% |

**p* < 0.05.

Figure 1. Prospective outcomes in the Standard Care Pathways of Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) according to number of previous episodes in subjects with bipolar disorder over one year.



depression during the study (Sachs et al., 2003, 2007). They were then allocated randomly to adjunctive antidepressant or placebo for up to 26 weeks. Depressive symptoms were then assessed with the continuous symptoms subscales for depression (SUM-D). This is part of the Clinical Monitoring Form, described and validated by Sachs et al. (2003).

We used different outcomes as validators of previously hypothesized staging models (Kapczinski et al., 2009), as well as empirical data (Berk et al., 2011a). As such, the outcomes available and tested here are on disability at baseline, the total number of days well and the longitudinal scores of depression, (hypo)mania, quality of life and functioning. Based on this, we constructed a priori statistical models, including available variables previously associated with outcome in bipolar disorder. A theoretical approach is preferred for the confirmation of hypothesis to other data-driven approaches as it avoids the problem of over-fitting (Babyak, 2004; Harrell et al., 1996). All models included baseline age (dichotomized at 65 years) and gender, low income (<\$US20,000 a year), not living with a partner, having less than any college education and being on

disability (except when it was the outcome). They also included the following clinical variables: current mood state (depressed, manic or euthymic), rapid cycling, current substance use or anxiety comorbidity, current smoking, childhood onset, baseline medications (lithium, other mood stabilizers, atypical antipsychotics, typical antipsychotics, benzodiazepines and antidepressants). Multiple dichotomic variables (<5, 5–10 and >10 episodes) were created to enter the number of previous episodes in the model. Finally, for the longitudinal analyses, an interaction term for study visit \times more than 10 relapses was entered. In all multivariate models, a single-level (i.e. not hierarchical) analysis was used. As such, in the results, the variable of interest is reported adjusted for all other variables in the model.

Statistical analysis

Different methods were used according to the outcome variable. For baseline disability we used logistic regression. For the total number of days well we used negative binomial regression (Elhai et al., 2008). For the open-label

Table 2. Multivariate effects of demographic and clinical variables on rating scale scores over one year in subjects with bipolar disorder in the Standard Care Pathways.

| | MADRS | YMRS | LIFE-RIFT | Q-LES |
|--|----------------|----------------|----------------|-----------------|
| Elderly | NS | NS | NS | NS |
| Female sex | NS | NS | -0.29 (0.12)* | NS |
| Low income | NS | NS | 0.26 (0.12)* | NS |
| Not being married | NS | 0.39 (0.18)* | NS | -1.23 (0.41)** |
| On disability | 1.84 (0.42)** | NS | 1.16 (0.16)*** | -2.31 (0.56)*** |
| Less than some college education | NS | NS | 0.35 (0.15)* | -1.25 (0.55)* |
| Smoking | 1.01 (0.34)* | NS | 0.28 (0.13)* | -1.81 (0.45)** |
| Any medical comorbidity | 0.95 (0.30)* | 0.69 (0.18)*** | 0.31 (0.16)** | -1.31 (0.40)** |
| Depression at baseline | 7.16 (0.32)** | NS | 2.18 (0.12)*** | -6.24(0.42)*** |
| Mania at baseline | NS | 2.87 (0.46)*** | NS | NS |
| Any anxiety disorder at baseline | 3.01 (0.33)** | 1.39 (0.20)*** | 0.80 (0.12)*** | -2.87(0.44)*** |
| Any substance use disorder at baseline | 1.11 (0.43)** | 1.15 (0.26)*** | NS | NS |
| Bipolar I disorder | NS | NS | | |
| Childhood onset | NS | 0.63 (0.21)** | NS | NS |
| Rapid cycling | 0.98 (0.34)** | 1.36 (0.20)*** | NS | -0.93 (0.46)* |
| Lifetime psychosis | NS | NS | NS | NS |
| Baseline medications | | | | |
| Lithium | NS | NS | NS | NS |
| Other mood stabilizers | NS | NS | NS | NS |
| Atypical antipsychotics | NS | -0.52 (0.20)** | 0.25 (0.13)* | NS |
| Typical antipsychotics | NS | -1.20 (0.57)* | NS | NS |
| Antidepressants | 1.10 (0.30)*** | NS | 0.37 (0.11)** | -1.63(0.40)*** |
| Benzodiazepines | 1.55 (0.35)*** | NS | 0.27 (0.13)* | -1.12 (0.47)* |
| 5–9 previous episodes | 1.23 (0.49)* | NS | NS | NS |
| 10 or more previous episodes | 2.66 (0.51)*** | 1.11 (0.30)*** | 0.97 (0.20)*** | -3.35(0.68)*** |

Scores shown are mean differences (S.E.) between groups. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

BDRS: Bipolar Depression Rating Scale; LIFE-RIFT: Longitudinal Interval Follow-up Evaluation – Range of Impairment Functioning Tool; MADRS: Montgomery-Asberg Rating Scale; Q-LES-Q: Quality of life Enjoyment and Satisfaction Questionnaire; YMRS: Young Mania Rating Scale.

longitudinal continuous data (MADRS, YMRS, RIFT, QLES), mixed effect models (Willett et al., 1998) were used. For the randomized pathway, we constructed a mixed effects model with the continuous SUM-D as the outcome. The model included time in treatment, group randomized and number of episodes as described above, plus two- and three-way interactions. Mixed effects models have many advantages over other methods of dealing with repeated

measures. They include the ability of using all available data on each subject and being unaffected by randomly missing data; time effects can be flexibly modeled, allowing for the use of parsimonious variance and correlation patterns (Gueorguieva and Krystal, 2004).

To be clear on the meaning of results of these models, ‘main effects’ of substantive predictors are differences between groups. These are variations on outcomes related

to the predictors. We also tested for interactions between predictors and time. When a predictor interacts with time, its impact on outcome varies with time period (Willett et al., 1998). Akaike values (Wagenmakers and Farrell, 2004) were used for model comparison, and likelihood ratio tests indicated a better adequacy for a random intercept and slope model with unstructured co-variances. Residuals were inspected for normality.

Results

Baseline characteristics of the STEP-BD cohort, as well as bivariate differences according to number of episodes, are presented in Table 1. Overall, 57% of the sample was comprised of women; the median age for the cohort at intake was 39 [interquartile range (IQR) 29–49] and the median age at first episode was 19 (IQR 15–26). At baseline, 10.3% of participants had fewer than five previous episodes, 38.1% had between 5 and 10 and 51.6% had more than 10 previous episodes.

The adjusted model for disability retained having more than 10 episodes as a predictor [odds ratio (OR) = 1.83; 95% confidence interval (CI) 1.15–2.89; $p = 0.010$]. Other clinical predictors were having a medical comorbidity (OR = 1.32; 95% CI 1.03–1.68; $p = 0.028$), bipolar I disorder (OR = 1.37; 95% CI 1.04–1.81; $p = 0.023$), smoking (OR = 1.41; 95% CI 1.11–1.80; $p = 0.005$) and being on an atypical antipsychotic (OR = 1.57; 95% CI 1.23–1.99; $p < 0.001$) or a mood stabilizing anticonvulsant at baseline (OR = 1.37; 95% CI 1.05–1.70; $p = 0.019$).

Having had more than 10 previous episodes (coefficient = -0.20; 95% CI -0.29 to -0.11; $p < 0.001$) and between 5 and 10 (coefficient -0.11; 95% CI -0.19 to -0.03; $p = 0.007$) were predictors of fewer days well; other predictors were depression at baseline (coefficient = -0.36; 95% CI -0.43 to -0.29; $p < 0.001$), any baseline anxiety (coefficient = -0.14; 95% CI -0.21 to -0.08; $p < 0.001$) or substance use disorders (coefficient = -0.08; 95% CI -0.17 to -0.00; $p = 0.050$), smoking (coefficient = -0.15; 95% CI -0.22 to -0.08; $p < 0.001$) and rapid cycling (coefficient = -0.09; 95% CI -0.17 to -0.02; $p = 0.011$), as well as use of benzodiazepines (coefficient = 0.08; 95% CI -0.15 to -0.01; $p = 0.024$).

Figure 1 depicts the course of clinical and functional variables according to the number of episodes over one year. Of note, significant effects indicated an improvement of all outcomes over time in all models ($p < 0.001$ for all). Table 2 illustrates between-group differences of all variables included in the models. The number of previous episodes was a significant predictor of each outcome, with no differences in time trajectories. For the MADRS, significant main effects for having more than 4 episodes ($F_{2015} = 6.43$, $p = 0.011$) and having more than 10 episodes ($F_{2456,548} = 30.21$, $p < 0.001$) were found, but not an interaction of having more than 10 episodes with time in treatment (i.e. the change trajectory) ($F_{1503,539} = 1.79$, $p = 0.182$).

Table 3. Baseline characteristics of subjects in the randomized pathway of STEP-BD according to treatment group

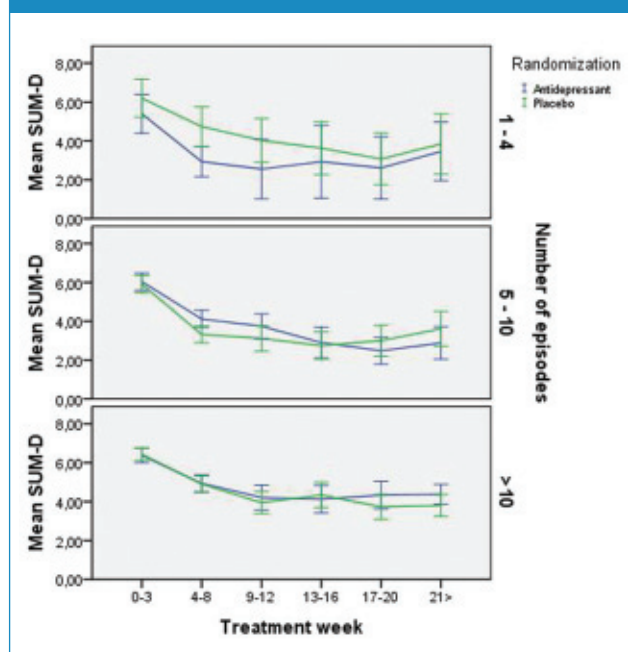
| | Mood stabilizer plus antidepressant (n = 179) | Mood stabilizer plus placebo (n = 187) |
|--|---|--|
| Age | 40.0 ± 22.3 | 40.0 ± 11.9 |
| Female sex | 57.6% | 56.1% |
| Low income (<\$US20,000 per year) | 42.1 | 48.6 |
| Not married | 64.9% | 69.3% |
| Less than some college education | 22.1% | 19.3% |
| Any anxiety disorder at baseline | 43.6% | 48.5% |
| Any substance use disorder at baseline | 16.7% | 15.7% |
| Bipolar I disorder | 68.6% | 67% |
| Rapid cycling | 27.2% | 32.5% |
| Baseline SUM-D | 6.2 ± 2.9 | 6.2 ± 3.1 |
| Baseline GAF | 56.0 ± 8.2 | 55.4 ± 7.8 |
| Fewer than 5 episodes | 8% | 10% |
| 5–9 previous episodes | 29% | 24% |
| 10 or more previous episodes | 63% | 66% |

GAF: Global Assessment of Functioning; SUM-D: continuous symptom subscale for depression.

Significant main effects for having more than more than 10 episodes ($F_{2628,049} = 12.19$, $p < 0.001$), but not an interaction of having more than 10 episodes with time in treatment ($F_{1550,101} = 0.48$, $p = 0.490$) were found on the YMRS. Main effects for more than 10 previous episodes ($F_{2087,197} = 22.10$, $p < 0.001$) were found for Q-LES scores, but not an interaction between number episodes and time ($F_{1136,594} = 0.07$, $p = 0.789$). For functioning, we found main effects for more than 10 episodes ($F_{2437,756} = 23.17$, $p < 0.001$), but not an interaction ($F_{1522,981} = 0.24$, $p = 0.627$).

In the model for the randomized pathway (Table 3 for baseline characteristics), the only significant effect was time in treatment ($F_{238,969} = 238.9723$, $p < 0.001$). Allocated group ($F_{308,072} = 0.38$, $p = 0.54$), having more than 4 ($F_{309,041} = 0.40$, $p = 0.53$) or 10 episodes ($F_{307,611} = 0.16$, $p = 0.69$) and all interactions were non-significant ($p > 0.05$ for all). Figure 2 displays evolution of SUM-D scores.

Figure 2. Evolution of depressive symptoms in the randomized pathway of Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) according to number of episodes and treatment group.



Discussion

The findings reported herein further confirm the utility of the concept of prognostic staging in bipolar disorder. Patients with multiple previous relapses had poor cross-sectional and prospective outcomes across the board. Functioning and quality of life were worse, disability more common, and symptoms more chronic and severe. Staging groups were different and remained different after one year of prospective guideline-informed treatment given at specialty clinics. This suggests that, on average, even after one year of intensive evidence-based treatment people with bipolar disorder in a late stage are unlikely to revert to previous functioning.

The risk associated with a chronically relapsing illness has been investigated before in prospective studies. Berk et al. (2011a) showed recently that those with more than five previous episodes had less favourable responses to treatment in an analysis of olanzapine treatment trials. In that study, acute and maintenance studies were pooled separately, and the focus was on syndromal outcome. Here, however, we took the approach of controlling for initial mood state and using mainly continuous rating scale scores as outcomes. This allowed us to see a parallel between symptom and functional outcomes, and their relation with number of previous episodes. Also, importantly, the emerging literature reveals a picture of incomplete recovery and frequent functional impairment in bipolar disorder.

Longitudinal assessment further shows that sub-threshold symptoms are the norm in inter-episode periods (Judd et al., 2005, 2008). In this report, we were further able to demonstrate that highly recurrent illness is associated with cross-sectional and longitudinal symptomatology and dysfunction independently of initial mood state.

One of the postulates of staging models is that early illness is amenable to interventions that are more effective and less harmful (Berk et al., 2007b; McGorry et al., 2010). A pertinent case in point would be adverse outcomes with psychotherapy for those with chronic illness (Berk and Parker, 2009). Two recent studies on cognitive-behaviour therapy and family psycho-education illustrate this. In those trials, late-stage was a meaningful predictor of poor outcomes (Reinares et al., 2010; Scott et al., 2006). We addressed this issue in the randomized pathway of the current study comparing antidepressants and placebo. While we were unable to find any significant differences, the STEP-BD dataset is not ideal to detect these differences as most patients were in one of the later stages. This would indicate that studies specifically designed to detect treatment effects in early stage illness are needed.

Not all longitudinal studies have found previous episodes as risk factors for worse outcomes. The definition of outcome may be germane to reconciling these differences. In a previous report, Perlis and colleagues (Perlis et al., 2006) failed to show convincingly that number of previous episodes predicted recurrences. They, however, used only a subset of STEP-BD participants who achieved recovery from an index episode. Aside from power issues, it is likely that the recovered subsample is different in terms of prognosis to the whole study sample. Interestingly, unlike in the current report, rapid cycling was a robust predictor of recurrence in that population. This suggests that predictors vary when different definitions of outcome are employed (Tedlow et al., 1998). While correlated, rapid cycling and a quantitative measure of previous relapses may represent two different processes. The former may be a superior predictor of recurrence, the latter a better proxy of cumulative morbidity.

To appreciate how repetitive illness episodes are translated into lower response to treatment and dysfunction the concepts of allostasis and neuroprogression are useful (Berk et al., 2011b; Kapczinski et al., 2008; McEwen, 2003; Magalhaes et al., 2012c; Moylan et al., 2012). The former predicts that the recurring mood episodes, substance abuse and medical comorbidity combine to produce a process of wear and tear that is associated with central and peripheral changes. Potentially mediated by various groups of biomarkers (Grande et al., 2012; Kapczinski et al., 2010, 2011), this 'cell endangerment' may create a vulnerability to further mood episodes. Although the direction of causality is uncertain at this moment, there are, indeed, many indications that functional, biochemical and even structural alterations accompany an increased number of

episodes (Post, 2010). This active process of neuroprogression would be responsible for the persistent functional impairment.

Regarding the confirmation of staging models, one of the limitations of this analysis is that even if the STEP-BD dataset is prospective, a longer period of observation is necessary to demonstrate the effect of transitioning between stages. This would, ideally, be demonstrated in inception cohorts with subjects at ‘ultra high’ risk and followed up well after transition to clinical illness. Another issue with the STEP-BD dataset is that the sample size falls with longer follow-up periods. That was the rationale to focus on one-year follow up.

Also of relevance, only relying on the number of previous episodes is a simplistic way of defining illness progression in bipolar disorder. As illness develops towards greater complexity and severity, several other features are likely to progress. Preliminary work, for instance, has demonstrated different brain derived neurotrophic factor (BDNF), cytokine and antioxidant profiles in early and late stages (Andreazza et al., 2009; Kauer-Sant’Anna et al., 2009). Showing longitudinal biological validation of staging models at this point is still needed (Berk et al., 2009). The number of previous episodes has, nevertheless, the virtue of being a simple, intuitive and clinically-relevant measure. As such, it is relevant to staging for researchers and clinicians (Berk et al., 2010b; McGorry et al., 2010). Furthermore, current diagnostic algorithms and treatment guidelines largely ignore measures of cumulative illness burden in their recommendations (McGorry et al., 2006). Although clinicians probably take previous courses into account in their decision-making, current clinical data are likely be weighted to the late stages. This is likely to obscure salient prognostic differences and complicate treatment selection (Berk et al., 2007a). Additional randomized data on specific interventions for different stages would constitute the next step for a more definitive validation of this construct.

Thus far, staging models have been used mostly as an argument for early intervention in psychiatry (Berk et al., 2010b). The experience with early interventions in schizophrenia has shown its value and demonstrated its possibilities (Bertelsen et al., 2008; Nordentoft et al., 2010). At this point, we would argue that late-stage bipolar disorder is also a syndrome necessitating differential attention. Similar results were recently reported by the Bipolar Collaborative Network (Post et al., 2010). In that study, patients very often used highly complex pharmacological regimens and a high number of previous episodes was associated with poor treatment response. With high rates of symptom chronicity, poor quality of life and persistent disability in spite of evidence-based treatment by trained clinicians, current treatments are clearly doing little to alleviate the burden of very substantial proportion of people with highly recurrent bipolar disorder.

As recently argued, psychiatry is likely to benefit from a notion of palliation (Berk et al., 2008, 2012). For some people, the goal of full symptomatic and functional recovery may be unrealistic. Nonetheless, this does not rule out a benefit from interventions focusing on functioning and quality of life. All in all, this study strengthens the argument for the need of developing specific interventions for highly recurrent and chronic bipolar disorder. Early intervention strategies are evidently meritorious as they may avert illness progression. However, many—even most—patients in tertiary facilities have late-stage bipolar disorder with limited benefit from current treatment strategies. These individuals clearly need specifically tailored late intervention strategies.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Declaration of interest

MB has received Grant/Research Support from the NIH, Cooperative Research Centre, Simons Autism Foundation, Cancer Council of Victoria, Stanley Medical Research Foundation, MBF, NHMRC, Beyond Blue, Rotary Health, Geelong Medical Research Foundation, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Organon, Novartis, Mayne Pharma and Servier, has been a speaker for Astra Zeneca, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Janssen Cilag, Lundbeck, Merck, Pfizer, Sanofi Synthelabo, Servier, Solvay and Wyeth, and served as a consultant to Astra Zeneca, Bristol Myers Squibb, Eli Lilly, Glaxo SmithKline, Janssen Cilag, Lundbeck Merck and Servier.

References

- Alonso J, Petukhova M, Vilagut G, Chatterji S, Heeringa S, Ustun TB, et al. (2011) Days out of role due to common physical and mental conditions: results from the WHO World Mental Health surveys. *Molecular Psychiatry* 16: 1234–1246.
- Andreazza AC, Kapczinski F, Kauer-Sant’Anna M, Walz JC, Bond DJ, Goncalves CA, et al. (2009) 3-Nitrotyrosine and glutathione antioxidant system in patients in the early and late stages of bipolar disorder. *Journal of Psychiatry & Neuroscience* 34: 236–271.
- Babak MA (2004) What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine* 66: 411–421.
- Berk M (2009) Neuroprogression: pathways to progressive brain changes in bipolar disorder. *International Journal of Neuropsychopharmacology* 12: 441–445.
- Berk M and Parker G (2009) The elephant on the couch: side-effects of psychotherapy. *Australian and New Zealand Journal of Psychiatry* 43: 787–794.
- Berk M, Conus P, Lucas N, Hallam K, Malhi GS, Dodd S, et al. (2007a) Setting the stage: from prodrome to treatment resistance in bipolar disorder. *Bipolar Disorders* 9: 671–678.
- Berk M, Hallam KT and McGorry PD (2007b) The potential utility of a staging model as a course specifier: A bipolar disorder perspective. *Journal of Affective Disorders* 100: 279–281.
- Berk M, Singh A and Kapczinski F (2008) When illness does not get better: do we need a palliative psychiatry? *Acta Neuropsychiatrica* 20: 165–166.

- Berk M, Malhi GS, Hallam K, Gama CS, Dodd S, Andreazza AC, et al. (2009) Early intervention in bipolar disorders: Clinical, biochemical and neuroimaging imperatives. *Journal of Affective Disorders* 114: 1–13.
- Berk M, Conus P, Kapczinski F, Andreazza AC, Yucel M, Wood SJ, et al. (2010a) From neuroprogression to neuroprotection: implications for clinical care. *Medical Journal of Australia* 193: S36–S40.
- Berk M, Hallam K, Malhi GS, Henry L, Hasty M, MacNeil C, et al. (2010b) Evidence and implications for early intervention in bipolar disorder. *Journal of Mental Health* 19: 113–126.
- Berk M, Brnabic A, Dodd S, Kelin K, Tohen M, Malhi GS, et al. (2011a) Does stage of illness impact treatment response in bipolar disorder? Empirical treatment data and their implication for the staging model and early intervention. *Bipolar Disorders* 13: 87–98.
- Berk M, Kapczinski F, Andreazza AC, Dean OM, Giorlando F, Maes M, et al. (2011b) Pathways underlying neuroprogression in bipolar disorder: focus on inflammation, oxidative stress and neurotrophic factors. *Neuroscience & Biobehavioral Reviews* 35: 804–817.
- Berk M, Berk L, Udina M, Moylan S, Stafford L, Hallam K, et al. (2012) Palliative models of care for later stages of mental disorder: maximizing recovery, maintaining hope, and building morale. *Australian and New Zealand Journal of Psychiatry* 46: 92–99.
- Bertelsen M, Jeppesen P, Petersen L, Thorup A, Ohlenschlaeger J, Le Quach P, et al. (2008) Five-year follow-up of a randomized multicenter trial of intensive early intervention vs standard treatment for patients with a first episode of psychotic illness: the OPUS trial. *Archives of General Psychiatry* 65: 762–771.
- Biewenga P, van der Velden J, Mo BW, Stalpers LJ, Schilthuis MS, van der Steeg JW, et al. (2009) Validation of existing prognostic models in patients with early-stage cervical cancer. *Gynecologic Oncology* 115: 277–284.
- Elhai JD, Calhoun PS and Ford JD (2008) Statistical procedures for analyzing mental health services data. *Psychiatry Research* 160: 129–136.
- Endicott J, Nee J, Harrison W and Blumenthal R (1993) Quality-of-Life Enjoyment and Satisfaction Questionnaire – a new measure. *Psychopharmacology Bulletin* 29: 321–326.
- Fava GA and Kellner R (1993) Staging - a neglected dimension in psychiatric classification. *Acta Psychiatrica Scandinavica* 87: 225–230.
- Francey SM, Nelson B, Thompson A, Parker AG, Kerr M, Macneil C, et al. (2010) Who needs antipsychotic medication in the earliest stages of psychosis? A reconsideration of benefits, risks, neurobiology and ethics in the era of early intervention. *Schizophrenia Research* 119: 1–10.
- Goldberg JF, Brooks JO, 3rd, Kurita K, Hoblyn JC, Ghaemi SN, Perlis RH, et al (2009) Depressive illness burden associated with complex polypharmacy in patients with bipolar disorder: findings from the STEP-BD. *Journal of Clinical Psychiatry* 70: 155–162.
- Grande I, Magalhaes PV, Kunz M, Vieta E and Kapczinski F (2012) Mediators of allostasis and systemic toxicity in bipolar disorder. *Physiology & Behavior* 106: 46–50.
- Gueorguieva R and Krystal JH (2004) Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry* 61: 310–317.
- Harrell FE, Lee KL and Mark DB (1996) Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine* 15: 361–387.
- Judd LL, Akiskal HS, Schettler PJ, Endicott J, Leon AC, Solomon DA, et al. (2005) Psychosocial disability in the course of bipolar I and II disorders – A prospective, comparative, longitudinal study. *Archives of General Psychiatry* 62: 1322–1330.
- Judd LL, Schettler PJ, Akiskal HS, Coryell W, Leon AC, Maser JD and Solomon DA (2008) Residual symptom recovery from major affective episodes in bipolar disorders and rapid episode relapse/recurrence. *Archives of General Psychiatry* 65: 386–394.
- Kapczinski F, Vieta E, Andreazza AC, Frey BN, Gomes FA, Tranontina J, et al. (2008) Allostatic load in bipolar disorder: Implications for pathophysiology and treatment. *Neuroscience and Biobehavioral Reviews* 32: 675–692.
- Kapczinski F, Dias VV, Kauer-Sant’Anna M, Frey BN, Grassi-Oliveira R, Colom F and Berk M (2009) Clinical implications of a staging model for bipolar disorders. *Expert Review of Neurotherapeutics* 9: 957–966.
- Kapczinski F, Dal-Pizzol F, Teixeira AL, Magalhaes PVS, Kauer-Sant’Anna M, Klamt F, et al. (2010) A systemic toxicity index developed to assess peripheral changes in mood episodes. *Molecular Psychiatry* 15: 784–786.
- Kapczinski F, Dal-Pizzol F, Teixeira AL, Magalhaes PV, Kauer-Sant’Anna M, Klamt F, et al. (2011) Peripheral biomarkers and illness activity in bipolar disorder. *Journal of Psychiatric Research* 45: 156–161.
- Kauer-Sant’Anna M, Kapczinski F, Andreazza AC, Bond DJ, Lam RW, Young LT and Yatham LN (2009) Brain-derived neurotrophic factor and inflammatory markers in patients with early- vs. late-stage bipolar disorder. *International Journal of Neuropsychopharmacology* 12: 447–458.
- Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott P and Andreasen NC (1987) The Longitudinal Interval Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* 44: 540–548.
- Kilbourne AM, Post EP, Nosssek A, Drill L, Cooley S and Bauer MS (2008) Improving medical and psychiatric outcomes among individuals with bipolar disorder: A randomized controlled trial. *Psychiatric Services* 59: 760–768.
- Kupfer DJ (2005) The increasing medical burden in bipolar disorder. *JAMA* 293: 2528–2530.
- McEwen BS (2003) Mood disorders and allostatic load. *Biological Psychiatry* 54: 200–207.
- McGorry PD (2007) Issues for DSM-V: Clinical staging: A heuristic pathway to valid nosology and safer, more effective treatment in psychiatry. *American Journal of Psychiatry* 164: 859–860.
- McGorry PD, Purcell R, Hickie IB, Yung AR, Pantelis C and Jackson HJ (2007) Clinical staging: a heuristic model for psychiatry and youth mental health. *Medical Journal of Australia* 187: S40–S42.
- McGorry PD, Hickie IB, Yung AR, Pantelis C and Jackson HJ (2006) Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Australian and New Zealand Journal of Psychiatry* 40: 616–622.
- McGorry PD, Nelson B, Goldstone S and Yung AR (2010) Clinical staging: a heuristic and practical strategy for new research and better health and social outcomes for psychotic and related mood disorders. *Canadian Journal of Psychiatry* 55: 486–497.
- MacNeil CA, Hallam K, Conus P, Henry L, Kader L and Berk M (2012) Are we missing opportunities for early intervention in bipolar disorder? *Expert Review of Neurotherapeutics* 12: 5–7.
- Magalhaes PV, Dean OM, Bush AI, Copolov DL, Weisinger D, Malhi G S, et al. (2012a) Systemic illness moderates the impact of N-acetyl cysteine in bipolar disorder. *Progress in Neuropsychopharmacology and Biological Psychiatry* 37: 132–135.
- Magalhaes PV, Kapczinski F, Nierenberg AA, Deckersbach T, Weisinger D, Dodd S and Berk M (2012b) Illness burden and medical comorbidity in the Systematic Treatment Enhancement Program for Bipolar Disorder. *Acta Psychiatrica Scandinavica* 125: 303–308.
- Magalhaes PVS, Fries GR and Kapczinski F (2012c) Peripheral markers and the pathophysiology of bipolar disorder. *Revista De Psiquiatria Clinica* 39: 60–67.
- Montgomery SA and Asberg M (1979) New depression scale designed to be sensitive to change. *British Journal of Psychiatry* 134: 382–389.
- Moylan S, Maes M, Wray NR and Berk M (2012) The neuroprogressive nature of major depressive disorder: pathways to disease evolution and resistance, and therapeutic implications. *Molecular Psychiatry*.

- Epub ahead of print 24 April 2012. DOI: 10.1038/mp.2012.3310.1038/mp.2012.33.
- Nordentoft M, Øhlenschlaeger J, Thorup A, Petersen L, Jeppesen P and Bertelsen M (2010) Deinstitutionalization revisited: a 5-year follow-up of a randomized clinical trial of hospital-based rehabilitation versus specialized assertive intervention (OPUS) versus standard treatment for patients with first-episode schizophrenia spectrum disorders. *Psychological Medicine* 40: 1619–1626.
- Otto MW, Simon NM, Wisniewski SR, Miklowitz DJ, Kogan JN, Reilly-Harrington NA, et al. (2006) Prospective 12-month course of bipolar disorder in out-patients with and without comorbid anxiety disorders. *British Journal of Psychiatry* 189: 20–25.
- Perlis RH, Ostacher MJ, Patel JK, Marangell LB, Zhang H, Wisniewski SR, et al. (2006) Predictors of recurrence in bipolar disorder: primary outcomes from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *American Journal of Psychiatry* 163: 217–224.
- Perlis RH, Dennehy EB, Miklowitz DJ, Delbello MP, Ostacher M, Calabrese JR, et al. (2009) Retrospective age at onset of bipolar disorder and outcome during two-year follow-up: results from the STEP-BD study. *Bipolar Disorders* 11: 391–400.
- Perlis RH, Ostacher MJ, Miklowitz DJ, Hay A, Nierenberg AA, Thase ME and Sachs GS (2010) Clinical features associated with poor pharmacologic adherence in bipolar disorder: Results from the STEP-BD study. *Journal of Clinical Psychiatry* 71: 296–303.
- Post RM (2010) Mechanisms of illness progression in the recurrent affective disorders. *Neurotoxicity Research* 18: 256–271.
- Post RM, Altshuler LL, Frye MA, Suppes T, Keck PE, McElroy SL, et al. (2010) Complexity of pharmacologic treatment required for sustained improvement in outpatients with bipolar disorder. *Journal of Clinical Psychiatry* 71: 1176–1186.
- Raballo A and Laro F (2009) Clinical staging: a new scenario for the treatment of psychosis. *Lancet* 374: 365–367.
- Reinases M, Colom F, Rosa AR, Bonnin CM, Franco C, Sole B, et al. (2010) The impact of staging bipolar disorder on treatment outcome of family psychoeducation. *Journal of Affective Disorders* 123: 81–86.
- Sachs GS, Guille C and McMurrich SL (2002) A clinical monitoring form for mood disorders. *Bipolar Disorders* 4: 323–327.
- Sachs GS, Thase ME, Otto MW, Bauer M, Miklowitz D, Wisniewski SR, et al. (2003) Rationale, design, and methods of the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biological Psychiatry* 53: 1028–1042.
- Sachs GS, Nierenberg AA, Calabrese JR, Marangell LB, Wisniewski SR, Gyulai L, et al. (2007) Effectiveness of adjunctive antidepressant treatment for bipolar depression. *New England Journal of Medicine* 356: 1711–1722.
- Scott J, Paykel E, Morriss R, Bentall R, Kinderman P, Johnson T, et al. (2006) Cognitive-behavioural therapy for severe and recurrent bipolar disorders - Randomised controlled trial. *British Journal of Psychiatry* 188: 313–320.
- Sheehan DV, Janavs J, Baker R, Harnett-Sheehan K, Knapp E, Sheehan M, et al. (1998) MINI – Mini International Neuropsychiatric Interview – English Version 5.0.0 – DSM-IV. *Journal of Clinical Psychiatry* 59: 34–57.
- Solomon DA, Leon AC, Coryell WH, Endicott J, Li C, Fiedorowicz JG, et al. (2010) Longitudinal course of bipolar I disorder: duration of mood episodes. *Archives of General Psychiatry* 67: 339–347.
- Tedlow J, Fava M, Uebelacker L, Nierenberg AA, Alpert JE and Rosenbaum J. (1998) Outcome definitions and predictors in depression. *Psychotherapy and Psychosomatics* 67: 266–270.
- Vieta E, Reinases M and Rosa AR (2011) Staging bipolar disorder. *Neurotoxicity Research* 19: 279–285
- Wagenmakers EJ and Farrell S (2004) AIC model selection using Akaike weights. *Psychonomic Bulletin and Review* 11: 192–196.
- Willett JB, Singer JD and Martin NC (1998) The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology* 10: 395–426.
- Young RC, Biggs JT, Ziegler VE and Meyer DA (1978) Rating-scale for mania – reliability, validity and sensitivity. *British Journal of Psychiatry* 133: 429–435.