

The Psychiatric Inclusion and Exclusion Criteria in Placebo-Controlled Monotherapy Trials of Bipolar Depression: An Analysis of Studies of the Past 20 Years

Mark Zimmerman^{1,2} · Carolina Guzman Holst¹ · Heather L. Clark¹ · Matthew Multach¹ · Emily Walsh¹ · Lia K. Rosenstein¹ · Douglas Gazarian¹

Published online: 19 August 2016
© Springer International Publishing Switzerland 2016

Abstract

Background Concerns about the generalizability of pharmacotherapy efficacy trials to “real-world” patients have been raised for more than 40 years. Almost all of this literature has focused on treatment studies of major depressive disorder (MDD).

Objective The aim of the study was to review the psychiatric inclusion and exclusion criteria used in placebo-controlled trials that assessed the efficacy of medications for bipolar depression (bipolar disorder efficacy trials [BDETs]) and compare the criteria used in BDETs with those used in efficacy trials of antidepressants to treat MDD (antidepressant efficacy trials [AETs]).

Methods We searched the MEDLINE, Embase, and PsycINFO databases for articles published from January 1995 through December 2014. We identified 170 placebo-controlled AETs and 22 BDETs published during these 20 years. Two of the authors independently reviewed each article and completed a pre-specified information extraction form listing the psychiatric inclusion and exclusion criteria used in the study.

Results Six inclusion/exclusion criteria were used in at least half of the BDETs: minimum severity on a depression symptom severity scale, significant suicidal ideation, diagnosis of alcohol or drug use disorder, presence of a comorbid nondepressive, nonsubstance use Axis I disorder, current episode of depression being too long, and absence of current manic symptoms. BDETs were significantly less likely than AETs to exclude patients with a history of psychotic features/disorders, borderline personality disorder, and post-traumatic stress disorder and more likely to exclude individuals who scored too low on the first item of the Hamilton Depression Rating Scale. Nearly two-thirds of the BDETs placed an upper limit on the duration of the current depressive episode, three times higher than the rate in the AETs. There was no difference on other variables between the AETs and BDETs.

Conclusions Similar to treatment studies of nonbipolar MDD, the treatment studies of bipolar depression frequently excluded patients with comorbid psychiatric and substance use disorders and insufficient severity of depressive symptoms as rated on standardized scales. These findings indicate that concerns about the generalizability of data from trials of recently approved medications for the treatment of bipolar depression are as relevant as the concerns that have been raised about studies of antidepressants for nonbipolar depression.

✉ Mark Zimmerman
mzimmerman@lifespan.org

¹ Department of Psychiatry, Rhode Island Hospital, Providence, RI, USA

² Department of Psychiatry and Human Behavior, Brown Medical School, Rhode Island Hospital, 146 West River Street, Providence, RI 02904, USA

Key Points

Concerns about the generalizability of pharmacotherapy efficacy trials to “real-world” patients have been raised for more than 40 years. However, almost all of this literature has focused on treatment studies of major depressive disorder (MDD).

We conducted the first review of placebo-controlled bipolar disorder efficacy trials (BDETs) and compared the psychiatric inclusion and exclusion criteria used in BDETs with those used in antidepressant efficacy trials (AETs) of MDD.

Similar to treatment studies of nonbipolar MDD, the treatment studies of bipolar depression frequently excluded patients with comorbid psychiatric and substance use disorders and insufficient severity of depressive symptoms as rated on standardized scales. These findings indicate that concerns about the generalizability of data from trials of recently approved medications for the treatment of bipolar depression are as relevant as the concerns that have been raised about studies of antidepressants for nonbipolar depression.

1 Introduction

Concerns about the generalizability of antidepressant efficacy trials (AETs) to “real-world” patients have been raised for more than 40 years [1–11]. An early line of inquiry compared the demographic, clinical, and treatment response characteristics of subjects who volunteered for efficacy treatment trials in response to advertisements (i.e., symptomatic volunteers) with patients who were referred for treatment in a more traditional manner. These studies generally found that the two groups were similar, and this was interpreted as support for the generalizability of efficacy trials [1–5]. However, other studies examined how many patients applying for an acute-phase efficacy trial were accepted into the trial and found low acceptance rates [6, 7]. In addition, more recently, the representativeness of samples treated in AETs was most directly examined by applying the inclusion and exclusion criteria to clinical samples. Several studies have found that most depressed outpatients treated in clinical practice would not qualify for an AET [8–11]. Almost all of this literature has focused on treatment studies of major depressive disorder (MDD).

During the past 15 years, greater attention has been given to the treatment of bipolar depression. For the first time, medications have received regulatory approval as monotherapy for the acute-phase treatment of bipolar depression without also being approved for the treatment of MDD. Accordingly, a sufficient literature of placebo-controlled studies has accumulated to warrant an examination of the psychiatric inclusion and exclusion criteria of trials that have assessed the efficacy of medications for bipolar depression (bipolar disorder efficacy trials [BDETs]).

We previously conducted a limited review of the psychiatric inclusion and exclusion criteria used in AETs by examining the criteria of 39 AETs published between 1994 and 2000 in five journals [12]. More recently, we conducted a comprehensive review of 170 placebo-controlled AETs published during the past 20 years [13]. We hypothesized that the concerns discussed a decade earlier would result in a broadening of the inclusion/exclusion criteria, thereby enhancing the generalizability of the data from AETs. Our hypothesis was not confirmed. Rather, an investigation of the inclusion/exclusion criteria of studies published during the past 5 years compared with those of studies published during the prior 15 years found that AETs have become more restrictive in the criteria used to select patients into the trials. The more recently published studies were significantly more likely to exclude depressed patients with any comorbid psychiatric disorder, more likely to require a minimum symptom duration longer than the *Diagnostic and Statistical Manual of Mental Disorders* fifth edition (DSM-5) 2-week threshold, and the cut-off score for inclusion in the study on measures of the severity of depressive symptoms was significantly higher in the more recent studies.

The goal of the present study was to conduct a similar review of placebo-controlled BDETs published during the past 20 years and to compare the criteria used in BDETs with those used in AETs.

2 Methods

We conducted a search of the MEDLINE (via PubMed), Embase (via Ovid), and PsycINFO (via EBSCO host) databases for articles published from January 1995 through December 2014. The search was limited to this timeframe to be consistent with our recently published review of the inclusion and exclusion criteria of AETs [13]. We used the search terms ‘depression’ or ‘depressive’ or ‘bipolar’ and ‘placebo’ and only included articles published in English. We also examined the reference lists of meta-analyses of AETs and BDETs and the studies identified from our literature review.

We excluded trials that focused on refractory depression; chronic depression; psychotic, atypical, or melancholic subtypes of depression; depressed patients with particular symptoms such as anxious features; inpatient samples; or samples that were limited to patients with a particular comorbid condition such as alcoholism, anxiety disorder, or medical illness. We excluded these studies from our analysis because, by definition, they focused on limited groups of depressed patients and this would bias our findings toward suggesting AETs and BDETs are poorly generalizable.

We only included trials focused on patients meeting DSM criteria for a current major depressive episode, and therefore did not include trials that were based on an admixture of patients with major depression, dysthymic disorder, cyclothymia, mixed bipolar disorder, and minor depression. Trials resulting in multiple publications based on the same sample (and the same set of inclusion/exclusion criteria) were included only once. We did not include trials of intravenous or injectable forms of medication or medication combinations or augmentation strategies. We included trials whether or not the medication had received regulatory approval for the treatment of depression.

Two of the authors independently reviewed each article and completed a pre-specified information extraction form listing the psychiatric inclusion and exclusion criteria used in the study. The reviewers met, compared the results of their data abstraction, and resolved discrepancies.

We identified 170 placebo-controlled AETs [13] and 22 BDETs published during the 20 years through December 2014 [14–30]. As shown in Table 1, three of the publications on bipolar disorder trials included the results of multiple trials [17, 20, 26]. In our tally of inclusion and exclusion criteria, we counted each trial separately. We compared the groups with the Chi square statistic or with Fisher's exact test if the expected value in any cell of a 2×2 table was <5 .

3 Results

3.1 Frequency of Psychiatric Inclusion/Exclusion Criteria used in Placebo-Controlled Monotherapy Treatment Studies of Bipolar Depression

Table 1 lists the inclusion and exclusion criteria used in each of the 22 BDETs, and Table 2 summarizes this information. Across all 22 BDETs, six inclusion/exclusion criteria were used in at least half of the studies: minimum severity on a depression symptom severity scale (100.0 %), significant suicidal ideation (81.8 %), diagnosis of alcohol

or drug abuse or dependence (77.3 %), presence of a comorbid nondepressive, nonsubstance use Axis I disorder (50.0 %), current episode of depression being too long (63.6 %), and current manic symptoms as reflected by scoring above a cut-off value on the Young Mania Rating Scale (77.3 %) [31].

The definition of some exclusion criteria varied between the studies. All 22 BDETs excluded patients whose symptom severity was too low. The most commonly used symptom severity measures were the 17-item Hamilton Rating Scale for Depression (HAMD) [32] (77.3 %, $n = 17$) and the Montgomery-Åsberg Depression Rating Scale (MADRS) [33] (27.3 %, $n = 6$). One study used both the HAMD and the MADRS [17]. Three different cut-offs on the HAMD were used for inclusion (cut-off of 16 in one study [18]; cut-off of 18 in nine studies [14, 15, 17, 26, 28]; cut-off of 20 in seven studies [16, 20, 22, 24, 25, 30]). Three different MADRS severity scores were used for study inclusion, the most common being a score of 20 (four studies [21, 23, 27, 29]).

Most studies excluded patients with current or recent substance use disorders (77.3 %, $n = 17$), though whether patients were excluded because of a history of drug/alcohol abuse, dependence, or either varied, as did the time period during which the patients could not have had substance use problems (current, past 3 months, past 6 months, or 1 year).

The range of nondepressive psychiatric pathology used as the basis for exclusion also varied among studies. Some listed the specific disorders that were the basis for exclusion, some listed diagnostic categories, and some indicated that patients with any comorbid diagnosis were excluded. The frequency of specific disorders used as exclusion criteria is presented in Table 3, which includes only disorders that were explicitly cited as exclusionary. For example, because no study specifically listed somatoform or impulse control disorders as exclusionary, these are not listed in the table. For studies that excluded patients with any comorbid disorder, we counted this as an exclusion for each disorder listed in the table. Some studies excluded patients with a limited number of disorders such as eating disorders [15, 17, 26], obsessive-compulsive disorder [15, 17, 26], or panic disorder [15, 17]. We did not count studies that limited exclusions to patients with a primary diagnosis of a nondepressive psychiatric disorder.

Nearly two-thirds of the 22 studies excluded patients based on the duration of the depressive episode. Exclusion because of an episode being too long (i.e., exclusion of patients with chronic depression) was more common than exclusion due to an episode being too short (i.e., a minimum episode duration requirement that is greater than the DSM-IV and DSM-5 requirements of 2 weeks). Studies excluded patients whose episodes were longer in duration

Table 1 Psychiatric inclusion/exclusion criteria in 22 bipolar depression efficacy trials

Study	Inclusion criteria	Exclusion criteria
Amsterdam and Shults [14]	Age ≥ 18 years BPI or II 17-item HAMD ≥ 18	Current alcohol or substance abuse Alcohol or substance dependence in past 3 months Dementia
Calabrese et al. [15] Same as Calabrese et al. [17] study 1, but slightly different description	Age ≥ 18 years BPI Episode duration ≤ 12 months 17-item HAMD ≥ 18 Two mood episodes in past 10 years, at least one manic or mixed	Rapid cycling Panic disorder, OCD, social phobia, or bulimia nervosa in past 12 months Substance dependence in past year or substance abuse in past month Actively suicidal
Calabrese et al. [16]	Age ≥ 18 and ≤ 65 years BPI or II 17-item HAMD ≥ 20 HAMD item 1 ≥ 2 Episode duration ≥ 1 month and ≤ 12 months	Axis I disorder as primary focus of treatment in past 6 months Substance dependence or substance use in past 12 months Serious suicidal or homicidal risk YMRS >12
Calabrese et al. [17] study 1	BPI 17-item HAMD ≥ 18 Episode duration ≤ 12 months Two mood episodes in past 10 years, at least one manic or mixed	Actively suicidal Rapid cycling Panic disorder, OCD, social phobia, or bulimia nervosa in past 12 months
Calabrese et al. [17] study 2	BPI or II 17-item HAMD ≥ 18 Episode duration ≤ 12 months BPI patients: two mood episodes in past 10 years, at least one manic or mixed BPII patients: one depressive and two hypomanic episodes in past 10 years	Actively suicidal Rapid cycling Panic disorder, OCD, social phobia, or bulimia nervosa in past 12 months
Calabrese et al. [17] study 3	BPI 17-item HAMD ≥ 18 Episode duration ≤ 12 months Two mood episodes in past 5 years, at least one manic or mixed MADRS ≥ 16	Actively suicidal Rapid cycling Panic disorder, OCD, social phobia, or bulimia nervosa in past 12 months SADS MRS ≥ 10
Calabrese et al. [17] study 4	BPII 17-item HAMD ≥ 18 HAMD item 1 or item 7 ≥ 3 Episode duration ≥ 8 weeks	Actively suicidal Rapid cycling Panic disorder, OCD, social phobia, or bulimia nervosa in past 12 months SADS MRS ≥ 10
Calabrese et al. [17] study 5	BPI 17-item HAMD ≥ 18 HAMD item 1 or item 7 ≥ 3 History of hospitalization or incarceration for mood episode Episode duration ≥ 8 weeks	Actively suicidal Rapid cycling Panic disorder, OCD, social phobia, or bulimia nervosa in past 12 months SADS MRS ≥ 10
Davis et al. [18]	BPI 17-item HAMD ≥ 16	Active Axis I disorder other than BPI disorder Borderline or antisocial personality disorder Significant suicidality Substance use disorder in past 3 months

Table 1 continued

Study	Inclusion criteria	Exclusion criteria
Ghaemi et al. [19]	BPI, II, or not otherwise specified MADRS ≥ 18	Current psychosis Severe suicidality Current substance abuse SADS MRS ≥ 12
Lombardo et al. [20] study 1	Age ≥ 18 years BPI Episode duration < 6 months 17-item HAMD ≥ 20	Schizophrenia, schizoaffective, schizophreniform, delusion disorder, or psychotic disorder not otherwise specified Psychotic features Ultrafast rapid cycling (eight or more mood episodes in past year) YMRS > 16 Alcohol or substance abuse in past 3 months Significant self-injurious/suicidal or violent behavior/homicidal risk
Lombardo et al. [20] study 2	Age ≥ 18 years BPI Episode duration < 6 months 17-item HAMD ≥ 20	Schizophrenia, schizoaffective, schizophreniform, delusion disorder, or psychotic disorder not otherwise specified Psychotic features Ultrafast rapid cycling (eight or more mood episodes in past year) YMRS > 16 Alcohol or substance abuse in past 3 months Significant self-injurious/suicidal or violent behavior/homicidal risk
Loebel et al. [21]	Age ≥ 18 and ≤ 75 years BPI Episode duration ≥ 1 and < 12 months MADRS ≥ 20	Psychotic features MADRS item 10 ≥ 4 Imminent risk of suicide or injury to self, others, or property YMRS > 12
McElroy et al. [22]	Age ≥ 18 years BPI or II 17-item HAMD ≥ 20 HAMD item 1 ≥ 2 Episode duration ≥ 4 weeks and ≤ 12 months	Axis I disorder as primary focus of treatment in past 6 months Substance dependence or substance use in past 12 months More than eight mood episodes in past year HAMD suicide item (item 3) ≥ 3 YMRS > 12
Muzina et al. [23]	Age ≥ 18 and ≤ 70 BPI or II MADRS ≥ 20	Significant suicide risk Alcohol, cocaine, or cannabis dependence in past 3 months Cocaine, hallucinogen, opiate, crystal meth, or MDMA abuse in past 3 months YMRS ≥ 12
Suppes et al. [24]	Age ≥ 18 and ≤ 65 years BPI or II Episode duration ≥ 1 and ≤ 12 months 17-item HAMD ≥ 20 HAMD item 1 ≥ 2	Axis I disorder in past 6 months Current substance abuse Current serious suicidal or homicidal risk or HAMD item 3 ≥ 3 Suicide attempt in past 6 months YMRS > 12

Table 1 continued

Study	Inclusion criteria	Exclusion criteria
Thase et al. [25]	Age ≥ 18 and ≤ 65 years BPI or II Episode duration ≥ 1 and ≤ 12 months 17-item HAMD ≥ 20 HAMD item 1 ≥ 2	Axis I disorder as primary focus of treatment in past 6 months Substance dependence or substance use in past 12 months Serious suicidal or homicidal risk YMRS >12
Thase et al. [26] study 1	Age ≥ 18 and ≤ 65 years BPI Episode duration ≤ 2 years 17-item HAMD ≥ 18 HAMD item 1 ≥ 2 At least one prior depressive episode	Primary psychiatric disorder other than BPI First depressive episode after age 55 years Six or more mood episodes in past year Cognitive or psychotic disorder Psychotic features Borderline or antisocial personality disorder Substance abuse in past 3 months or substance dependence in past 6 months OCD, bulimia nervosa, or ADHD in past 3 months Significant risk of suicide YMRS >12
Thase et al. [26] study 2	Age ≥ 18 and ≤ 65 years BPI Episode duration ≤ 2 years 17-item HAMD ≥ 18 HAMD item 1 ≥ 2 At least one prior depressive episode	Primary psychiatric disorder other than BPI First depressive episode after age 55 years Six or more mood episodes in past year Cognitive or psychotic disorder Psychotic features Borderline or antisocial personality disorder Substance abuse in past 3 months or substance dependence in past 6 months OCD, bulimia nervosa, or ADHD in past 3 months Significant risk of suicide YMRS >12
Tohen et al. [27]	Age ≥ 18 years BPI Prior manic/mixed episode requiring treatment MADRS ≥ 20	Alcohol or substance dependence in past 3 months Suicidal behavior within past 3 months
Tohen et al. [28]	Age ≥ 18 and < 65 years BPI Prior manic/mixed episode in past 6 years Episode duration ≤ 90 days 17-item HAMD ≥ 18	Recent substance dependence A history of serious psychiatric illness other than bipolar disorder Current rapid-cycling mood disturbance YMRS ≥ 8
Wang et al. [29]	Age ≥ 18 and ≤ 60 years BPI MADRS ≥ 20 CGI-S ≥ 4	Substance dependence in past 6 months Mental retardation YMRS ≥ 15
Young et al. [30]	Age ≥ 18 and ≤ 65 years BPI or II Episode duration ≥ 4 weeks and ≤ 12 months 17-item HAMD ≥ 20 HAMD item 1 ≥ 2	Active Axis I disorders requiring treatment within 6 months Substance dependence or abuse Current serious suicidal or homicidal risk YMRS >12

ADHD attention-deficit/hyperactivity disorder, *BPI/II* bipolar I/II, *CGI-S* Clinical Global Impression Severity scale, *HAMD* Hamilton Depression Rating Scale, *MADRS* Montgomery-Åsberg Depression Rating Scale, *MMDA* methoxymethylenedioxyamphetamine, *OCD* obsessive-compulsive disorder, *SADS MRS* Schedule for Affective Disorders and Schizophrenia Mania Rating Scale, *YMRS* Young Mania Rating Scale

Table 2 Commonly used psychiatric inclusion/exclusion criteria in 170 antidepressant efficacy trials [13] and 22 bipolar depression efficacy trials

Exclusion criterion	AETs (<i>n</i> = 170)	BDETs (<i>n</i> = 22)	χ^2	<i>p</i> value
Severity scale score below cut-off	170 (100.0)	22 (100.0)	Fisher	NS
Severity scale score above cut-off	14 (8.2)	0 (0.0)	Fisher	NS
Psychotic disorder/current psychotic features	143 (84.1)	10 (45.5)	Fisher	<0.01
Substance abuse/dependence	137 (80.6)	17 (77.3)	Fisher	NS
Significant suicidal ideation	128 (75.3)	18 (81.8)	0.46	NS
History of suicide attempt(s)	35 (20.6)	4 (18.2)	Fisher	NS
Significant homicidal ideation/violence risk	28 (16.5)	7 (31.8)	Fisher	NS
Other nondepressive/nonsubstance use disorders	92 (54.1)	11 (50.0)	0.13	NS
Episode duration too long	34 (20.0)	14 (63.6)	19.78	<0.001
Episode duration too short	81 (47.6)	8 (36.4)	1.00	NS
Borderline personality disorder	70 (41.2)	4 (18.2)	4.35	<0.05
Item 1 Hamilton Rating Scale below cut-off	33 (19.4)	9 (40.9)	Fisher	<0.05
Any Axis I disorder	46 (27.1)	3 (13.6)	1.85	NS
Presence of manic symptoms	0 (0.0)	16 (77.3)	134.9	<0.001

Data are presented as *n* (%) unless otherwise indicated

AET antidepressant efficacy trial, BDET bipolar depression efficacy trials, NS not significant

Table 3 Number of antidepressant efficacy trials [13] and bipolar depression efficacy trials excluding patients with different nondepressive disorders

Diagnosis	AETs (<i>n</i> = 170)	BDETs (22 studies)	χ^2	<i>p</i> value
None	17 (10.0)	1 (4.5)	Fisher	NS
Any Axis I disorder	46 (27.1)	3 (13.6)	1.85	NS
Panic disorder	65 (38.2)	8 (36.4)	0.03	NS
Generalized anxiety disorder	50 (29.4)	3 (13.6)	2.43	NS
Social anxiety disorder	48 (28.2)	8 (36.4)	0.62	NS
Obsessive–compulsive disorder	80 (47.1)	11 (50.0)	0.07	NS
Posttraumatic stress disorder	75 (44.1)	4 (18.2)	5.41	<0.05
Specific phobia	45 (26.5)	3 (13.6)	1.71	NS
Alcohol abuse	129 (75.9)	13 (59.1)	2.85	NS
Alcohol dependence	109 (64.1)	14 (63.6)	0.00	NS
Drug abuse	130 (76.5)	14 (63.6)	1.71	NS
Drug dependence	107 (62.9)	14 (63.6)	0.00	NS
Anorexia nervosa	66 (38.8)	4 (18.2)	3.58	NS
Bulimia nervosa	66 (38.8)	11 (50.0)	1.01	NS
Any personality disorder	60 (35.3)	0 (0.0)	11.29	<0.001
Borderline personality disorder	70 (41.2)	4 (18.2)	4.35	<0.05
Antisocial personality disorder	68 (40.0)	4 (18.2)	3.96	<0.05
Schizotypal personality disorder	63 (37.1)	0 (0.0)	12.14	<0.001

Data are presented as *n* (%) unless otherwise indicated

AET antidepressant efficacy trial, BDET bipolar depression efficacy trials, NS not significant

than 3 months (*n* = 1) [28], 6 months (*n* = 2) [20], 1 year (*n* = 9) [15–17, 21, 22, 24, 25, 30], or 2 years (*n* = 2) [26]. Eight studies required the episode duration to be greater than the 2 weeks required by the DSM-IV and DSM-5 to make the diagnosis. Of these eight studies, six required a minimum symptom duration of 1 month [16, 21, 22, 24, 25, 30] and two required a minimum duration of 2 months [17].

3.2 Comparison of the Psychiatric Exclusion Criteria used in Antidepressant Efficacy Trials (AETs) and Bipolar Disorder Efficacy Trials (BDETs)

BDETs were significantly less likely than AETs to exclude patients with current or a history of psychotic features/disorders (Table 2). Nearly two-thirds of the BDETs

placed an upper limit on the duration of the current depressive episode, more than three times higher than the 20 % rate in the AETs. There was no difference between the AETs and BDETs in exclusion due to total score on a depression symptom scale, history of suicide attempts, current suicidal ideation, or the depressive episode being too short, though the BDETs were significantly more likely to exclude patients who did not score high enough on the first item of the HAMD.

Approximately one-third of the AETs excluded patients with any personality disorder, whereas none of the BDETs used this as an exclusion criterion (Table 3). Post-traumatic stress disorder, and borderline, schizotypal, and antisocial personality disorder were significantly more often listed as exclusions in AETs. There were no significant differences between the AETs and BDETs in exclusion due to other anxiety disorders, substance use disorders, or eating disorders.

4 Discussion

The goals of the present study were to review the psychiatric inclusion/exclusion criteria used in BDETs published during the past 20 years and to compare the criteria used in BDETs with those of AETs. More than a decade ago, our clinical research group raised concerns about the generalizability of AETs and suggested that the majority of patients seen in routine clinical practice would not qualify for an AET [8]. Our findings were subsequently independently replicated by other clinical research groups [9–11]. We recently updated and expanded our initial review of AETs, and found that AETs continue to recruit a narrow range of depressed patients [13]. To the best of our knowledge, the present review is the first to examine the inclusion/exclusion criteria of monotherapy treatment trials of bipolar depression. To enhance comparability to the monotherapy AETs we previously reviewed, we focused on monotherapy studies of bipolar depression and did not include augmentation trials.

Our comparison of BDETs and AETs found more similarities than differences. All studies required a minimum level of severity on a measure of depressive symptoms, and almost all studies excluded patients with at least some comorbid psychiatric disorders. Most BDETs and AETs excluded patients with suicidal ideation at the time of screening for the treatment study, and most also excluded patients with a history of substance use disorders. The studies of bipolar disorder were significantly less likely to exclude patients with psychotic features, a potentially important difference because patients with bipolar I depression are more likely to have psychotic symptoms than patients with nonbipolar depression [34].

To diagnose a major depressive episode, DSM-5, like its recent predecessors, requires that the symptoms of depression be present for at least 2 weeks. Yet, a significant minority of AETs and BDETs required a minimum symptom duration of at least 1 month. The minimum symptom duration requirement for AETs was more common than the maximum symptom duration exclusion. This was reversed for the BDETs, and exclusion due to chronicity was more frequent than exclusion due to brief symptom duration. The reason for this difference between AETs and BDETs is not clear, as few studies justify the basis for the exclusion criteria used.

Subject selection in BDETs and AETs must balance the issues of internal and external validity. In AETs, the inclusion/exclusion criteria have narrowed over the past 5 years, thereby suggesting that AETs may be even less generalizable than they were previously (when concerns about their generalizability had already been raised) [13]. We were unable to similarly examine whether the inclusion/exclusion criteria of BDETs have changed during the past 20 years because an insufficient number of studies have been conducted to permit such an analysis. The percentage of BDETs published during the past 5 years was not significantly different than the percentage of AETs published during the same 5-year period (40.9 vs. 32.9 %, $\chi^2 = 0.74$, not significant).

A limitation of the present analysis is that it was based on published placebo-controlled studies. There is evidence of publication bias in AETs [35] and BDETs [36], and it is possible that the inclusion/exclusion criteria of published and unpublished studies systematically differ. However, recent reviews of the literature on agomelatine compared efficacy in all published and unpublished studies conducted on the medication and provided links to the unpublished studies [37, 38]. We downloaded the referenced unpublished trials from the internet and found that the inclusion/exclusion criteria of the published and unpublished trials were essentially the same. Similarly, a review of studies of lamotrigine found the criteria in published and unpublished studies were nearly identical [17].

5 Conclusions

This is the first study of the inclusion and exclusion criteria used in monotherapy treatment studies of bipolar depression. Similar to treatment studies of nonbipolar MDD, the treatment studies of bipolar depression frequently excluded patients with comorbid psychiatric and substance use disorders and insufficient severity of depressive symptoms as rated on standardized scales. Thus, concerns about the generalizability of data from trials of recently approved medications for the treatment of bipolar depression are as

relevant as the concerns that have been raised about studies of antidepressants for nonbipolar depression.

Compliance with Ethical Standards

Funding source No sources of funding were used to conduct this study or prepare this manuscript.

Conflicts of interest M Zimmerman has participated on advisory boards for F. Hoffmann-La Roche, Genentech, and Lundbeck; received research support from Azevan and Eli Lilly; and prepared educational material for Otsuka. Ms. Holst, Ms. Clark, Mr. Multach, Ms. Walsh, Ms. Rosenstein, and Mr. Gazarian do not have any conflicts.

References

- Brauzer B, Goldstein BJ. Symptomatic volunteers: another patient dimension for clinical trials. *J Clin Pharmacol.* 1973;13:89–98.
- Breckenridge JS, Zeiss AM, Breckenridge JN, Gallagher D, Thompson LW. Solicitations of elderly depressives for treatment outcome research: a comparison of referral sources. *J Consult Clin Psychol.* 1985;53:552–4.
- Miller CA, Hooper CL, Bakish D. A comparison of patients with major depressive disorder recruited through newspaper advertising versus consultation referrals for clinical drug trials. *Psychopharmacol Bull.* 1997;33:69–73.
- Rapaport MH, Zisook S, Frevert T, Seymour S, Kelsoe JR, Judd LJ. A comparison of descriptive variables for clinical patients and symptomatic volunteers with depressive disorders. *J Clin Psychopharmacol.* 1996;16:242–6.
- Covi L, Lipman R, McNair D, Czerlinsky T. Symptomatic volunteers in multicenter drug trials. *Neuropsychopharmacol.* 1979;3:521–33.
- Partonen T, Sihvo S, Lonnqvist JK. Patients excluded from antidepressant efficacy trial. *J Clin Psychiatry.* 1996;57:572–5.
- Keitner G, Posternak M, Ryan C. How many subjects with major depressive disorder meet eligibility requirements of an antidepressant efficacy trial? *J Clin Psychiatry.* 2003;64:1091–3.
- Zimmerman M, Mattia JJ, Posternak MA. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice. *Am J Psychiatry.* 2002;159:469–73.
- van der Lem R, van der Wee NJ, van Veen T, Zitman FG. The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychol Med.* 2011;41:1353–63.
- Wisniewski SR, Rush AJ, Nierenberg AA, Gaynes BN, Warden D, Luther JF, et al. Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR*D report. *Am J Psychiatry.* 2009;166:599–607.
- Zetin M, Hoepner CT. Relevance of exclusion criteria in antidepressant clinical trials: a replication study. *J Clin Psychopharmacol.* 2007;27:295–301.
- Zimmerman M, Chelminski I, Posternak M. Exclusion criteria used in antidepressant efficacy trials: consistency across studies and representativeness of samples included. *J Nerv Ment Dis.* 2004;192:87–94.
- Zimmerman M, Clark HL, Multach MD, Walsh E, Rosenstein LK, Gazarian D. Have treatment studies of depression become even less generalizable?: a review of the inclusion and exclusion criteria in placebo controlled antidepressant efficacy trials published during the past 20 years. *Mayo Clin Proceed.* 2015;90:1180–6.
- Amsterdam J, Shults J. Comparison of fluoxetine, olanzapine, and combined fluoxetine plus olanzapine initial therapy of bipolar type I and type II major depression—lack of manic induction. *J Affect Disord.* 2005;87:121–30.
- Calabrese J, Bowden C, Sachs G, Ascher J, Monaghan E, Rudd G. A double-blind placebo-controlled study of lamotrigine monotherapy in outpatients with bipolar I depression. Lamictal 602 Study Group. *J Clin Psychiatry.* 1999;60:79–88.
- Calabrese J, Keck P, Macfadden W, Minkwitz M, Ketter T, Weisler R, et al. A randomized, double-blind, placebo-controlled trial of quetiapine in the treatment of bipolar I or II depression. *Am J Psychiatry.* 2005;162:1351–60.
- Calabrese J, Huffman R, White R, Edwards S, Thompson T, Ascher J, et al. Lamotrigine in the acute treatment of bipolar depression: results of five double-blind, placebo-controlled clinical trials. *Bipolar Disord.* 2008;10:323–33.
- Davis L, Bartolucci A, Petty F. Divalproex in the treatment of bipolar depression: a placebo-controlled study. *J Affect Disord.* 2005;85:259–66.
- Ghaemi S, Gilmer W, Goldberg J, Zablotsky B, Kemp D, Kelley M, et al. Divalproex in the treatment of acute bipolar depression: a preliminary double-blind, randomized, placebo-controlled pilot study. *J Clin Psychiatry.* 2007;68:1840–4.
- Lombardo I, Sachs G, Kolluri S, Kremer C, Yang R. Two 6-week, randomized, double-blind, placebo-controlled studies of ziprasidone in outpatients with bipolar I depression: did baseline characteristics impact trial outcome? *J Clin Psychopharmacol.* 2012;32:470–8.
- Loebel A, Cucchiaro J, Silva R, Kroger H, Hsu J, Sarma K, et al. Lurasidone monotherapy in the treatment of bipolar I depression: a randomized, double-blind, placebo-controlled study. *Am J Psychiatry.* 2014;171:160–8.
- McElroy S, Weisler R, Chang W, Olausson B, Paulsson B, Brecher M, et al. A double-blind, placebo-controlled study of quetiapine and paroxetine as monotherapy in adults with bipolar depression (EMBOLDEN II). *J Clin Psychiatry.* 2010;71:163–74.
- Muzina D, Gao K, Kemp D, Khalife S, Ganocy S, Chan P, et al. Acute efficacy of divalproex sodium versus placebo in mood stabilizer-naïve bipolar I or II depression: a double-blind, randomized, placebo-controlled trial. *J Clin Psychiatry.* 2011;72:813–9.
- Suppes T, Datto C, Minkwitz M, Nordenhem A, Walker C, Darko D. Effectiveness of the extended release formulation of quetiapine as monotherapy for the treatment of acute bipolar depression. *J Affect Disord.* 2014;168:485–93.
- Thase M, Macfadden W, Weisler R, Chang W, Paulsson B, Khan A, et al. Efficacy of quetiapine monotherapy in bipolar I and II depression: a double-blind, placebo-controlled study (the BOLDER II study). *J Clin Psychopharmacol.* 2006;26:600–9.
- Thase M, Jonas A, Khan A, Bowden C, Wu X, McQuade R, et al. Aripiprazole monotherapy in nonpsychotic bipolar I depression: results of 2 randomized, placebo-controlled studies. *J Clin Psychopharmacol.* 2008;28:13–20.
- Tohen M, Vieta E, Calabrese J, Ketter T, Sachs G, Bowden C, et al. Efficacy of olanzapine and olanzapine-fluoxetine combination in the treatment of bipolar I depression. *Arch Gen Psychiatry.* 2003;60:1079–88.
- Tohen M, McDonnell D, Case M, Kanba S, Ha K, Fang Y, et al. Randomised, double-blind, placebo-controlled study of olanzapine in patients with bipolar I depression. *Br J Psychiatry.* 2012;201:376–82.
- Wang M, Tong J, Huang D, Zhu G, Liang G, Du H. Efficacy of olanzapine monotherapy for treatment of bipolar I depression: a randomized, double-blind, placebo controlled study. *Psychopharmacology.* 2014;231:2811–8.

30. Young A, McElroy S, Bauer M, Philips N, Chang W, Olausson B, et al. A double-blind, placebo-controlled study of quetiapine and lithium monotherapy in adults in the acute phase of bipolar depression (EMBOLDEN I). *J Clin Psychiatry*. 2010;71:150–62.
31. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry*. 1978;133:429–35.
32. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56–62.
33. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382–9.
34. Goes FS, Sadler B, Toolan J, Zamoiski RD, Mondimore FM, MacKinnon DF, et al. Psychotic features in bipolar and unipolar depression. *Bipolar Disord*. 2007;9:901–6.
35. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358:252–60.
36. Ghaemi SN. The failure to know what isn't known: negative publication bias with lamotrigine and a glimpse inside peer review. *Evid Based Ment Health*. 2009;12:65–8.
37. Koesters M, Guaiana G, Cipriani A, Becker T, Barbui C. Agomelatine efficacy and acceptability revisited: systematic review and meta-analysis of published and unpublished randomised trials. *Br J Psychiatry*. 2013;203:179–87.
38. Taylor D, Sparshatt A, Varma S, Olofinjana O. Antidepressant efficacy of agomelatine: meta-analysis of published and unpublished studies. *BMJ*. 2014;348:g1888.