Journal of Affective Disorders 223 (2017) 41–48

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

# Efficacy and safety of long-term antidepressant treatment for bipolar disorders – A meta-analysis of randomized controlled trials

Bangshan Liu[a,1], Yan Zhang[a,1], Han Fang[a], Jin Liu[a], Tiebang Liu[b], Lingjiang Li[a,b,*]

[a] Mental Health Institute, the Second Xiangya Hospital of Central South University, National Clinical Research Center for Mental Disorder, National Technology Institute of Psychiatry, Key Laboratory of Psychiatry and Mental Health of Hunan Province, Changsha 410011, China
[b] Shenzhen Kang Ning Hospital, Shenzhen, Guangdong, China

ARTICLE INFO

*Keywords:*
Bipolar disorder
Antidepressant
Long-term
Meta-analysis

ABSTRACT

*Objective:* Efficacy and safety of long-term use of antidepressants (AD) in bipolar disorder (BD) patients remains highly controversial. Here we performed a meta-analysis of randomized controlled trials (RCTs) exploring the efficacy and safety of long-term AD use in BD patients.

*Methods:* English-written literature published in peer-reviewed journal was systematically searched from Pubmed, EMBASE, CENTRAL, PsycINFO and Clinicaltrials.gov. Each database was searched from its first available time to August 31, 2016. Additional papers were searched from recent guidelines, expert consensus and systematic reviews by hand. RCTs exploring the efficacy and safety of long-term (≥ 4 m) antidepressant treatment for patients with bipolar disorder were eligible. Two authors (HF, JL) independently extracted the data. Risk ratio (RR), number needed to treat (NNT) and/or number needed to harm (NNH) for new depressive episodes and new manic/hypomanic episodes were calculated. Subgroup analyses were performed based on treatment regimen (AD monotherapy or combined with MS), types of antidepressants, funding source, bipolar subtypes and treatment duration.

*Results:* Eleven trials with 692 bipolar disorder patients were included in the meta-analysis. The risk of bias assessment demonstrated moderate bias risk. Antidepressants were superior to placebo in reducing new depressive episodes in bipolar disorders without increasing risk of new manic/hypomanic episodes either used as monotherapy or in combination with MS. Subgroup analyses revealed that greater benefit and lower risk may be achieved in BD II than in BD I. However, compared with MS monotherapy, AD monotherapy significantly increased the risk of affective switch with no improvement in prophylaxis of new depressive episodes.

*Conclusions:* Reduced new depressive episodes may be achieved by long-term AD treatment with no significantly increased risk of new manic/hypomanic episodes in BD, particularly in BD II. The elevated risk of affective switch of AD monotherapy compared with MS monotherapy may be contributed to the protective effect of MS in diminishing manic/hypomanic episodes. Further studies are needed to verify our findings.

## 1. Introduction

Depressive episodes in bipolar disorder (BD) usually exhibit higher prevalence, longer duration, more serious harm on social function and higher burden than manic/hypomanic episodes (Bopp et al., 2010; Judd et al., 2002; Miller et al., 2014; Michalak et al., 2008; Solomon et al., 2010), while the treatment of bipolar depression, especially long-term treatment, is much less studied and far less optimized in clinical practice than mania/hypomania (Grunze et al., 2013). Practice guidelines and expert consensus (Grunze et al., 2013; Goodwin, 2009; Yatham et al., 2013; Malhi et al., 2015; Pacchiarotti et al., 2013) recommend avoiding use of antidepressants (AD) for BD patients, only if the depressive episode is very severe and shows poor response to mood stabilizers (MS) or atypical antipsychotics (AP) monotherapy. And, even in the case of indispensable use of AD for acute treatment of bipolar depression, the use of AD is recommended to be limited in 6–8 weeks after full remission of depression (Yatham et al., 2013), much shorter than the recommended continuation and maintenance treatment duration in unipolar major depressive disorder (MDD)(Davidson, 2010). Nevertheless, these recommendations have never been backed up by explicit evidence.

Although discouraged from guidelines, the long-term AD use in BD

* Correspondence to: No. 139, Renmin Middle Road, Changsha, Hunan, China.
*E-mail address:* llj2920@163.com (L. Li).
[1] The authors contributed equally to the work.

http://dx.doi.org/10.1016/j.jad.2017.07.023
Received 11 April 2017; Received in revised form 7 June 2017; Accepted 8 July 2017
Available online 11 July 2017

patients is commonly seen in clinical practice, with as high as 40% of patients taking AD in the maintenance phase (Grande et al., 2013). This may be due to the inadequate effectiveness or poor tolerability of recommended treatment regimens, namely, the MSs and APs. However, the use of antidepressant in bipolar disorder (BD) treatment, especially long-term treatment, is highly contentious since two of most commonly concerned questions about this topic, namely, the effectiveness of long-term AD treatment for prophylaxis of new depressive episodes and the risk of new manic/hypomanic episodes inducement are exceedingly inconclusive due to the limited and inconsistent evidence from clinical trials (Pacchiarotti et al., 2013; Gitlin, 2012; Ghaemi, 2012).

A previous meta-analysis exploring the benefits and risk of long-term AD treatment for BD patients demonstrated that long-term adjunctive AD treatment had little protection for depression relapse while significantly increased risk of affective switch (Ghaemi et al., 2008). However, the results of this meta-analysis may be biased by repeated using of data for more than one time in multiple-arm trials. Besides, the authors didn't differentiate comparisons between AD and placebo from comparisons between AD and MS, which may result in confounding results since MS and placebo is considered explicitly to have different effects on depression relapse prevention and affective switch inducement. Moreover, several new studies with different findings have been published since the publication of the meta-analysis. Therefore, we conducted this updated meta-analysis of randomized controlled trials (RCTs) to address the efficacy and affective switch risk of long-term AD use in BD patients.

## 2. Materials and methods

### 2.1. Literature search

Clinical trials were searched from Pubmed, EMBASE, The Cochrane Central Register of Controlled Trials (CENTRAL), PsyINFO and Clinicaltrials.gov using "bipolar disorder", "manic-depressive disorder", "antidepressant or tricyclic OR tetracyclic OR serotonin reuptake inhibitor OR noradrenaline reuptake inhibitor OR monoamine oxidase inhibitor OR sertraline OR venlafaxine OR escitalopram OR citalopram OR paroxetine OR fluoxetine OR mirtazapine OR fluvoxamine OR bupropion OR wellbutrin OR duloxetine OR trazodone OR nefazodone OR reboxetine OR moclobemide OR mianserine OR amitriptyline OR chlorimipramine OR St John's wort OR hypericum OR imipramine OR doxepin OR maprotiline", "randomized controlled trials". Publication date was restricted from every database's first available time to 2016/08/31. Free text search and Mesh search were combined to improve the recall ratio. The references of included studies and guidelines were also systematically searched by hand. The detailed description of search strategies and results in each database was shown in Appendix.

### 2.2. Inclusion criteria

Studies included in this meta-analysis should meet the following criteria:

(1) Inclusion of patients diagnosed with Bipolar I, II or NOS type by Diagnostic and Statistical Manual of Mental Disorders (DSM-III, DSM-IV, DSM-IV-TR, DSM-5) or Bipolar Disorder by International Classification of Diseases (ICD-9, ICD-10) or research domain criteria (RDC) criteria. When both unipolar and bipolar depressed patients were included, the data of bipolar depressed patients should be reported separately.
(2) Allocation of patients should be based on randomization.
(3) The average total duration of AD treatment should be lasted for at least 4 months (we defined the "long-term" as ≥4 m based on the criteria used in a previous meta-analysis exploring the efficacy and safety of antidepressant for acute treatment (16 weeks) of bipolar disorder (Sidor and Macqueen, 2011)).
(4) Data about the treatment-emergent of new depressive episodes or manic/hypomanic episodes during long-term antidepressant treatment for patients who achieved clinical remission or response should be reported.

### 2.3. Exclusion criteria

(1) Naturalistic study or observational study design.
(2) Only data about efficacy and safety of acute AD treatment for bipolar depression was reported.
(3) Study protocols or inadequate data reporting (no available data for meta-analysis).
(4) Repeated reporting.

### 2.4. Data extraction

Data about the demographic information (gender, age), index episode (depression, mania or not specified), inclusion criteria (diagnosis, severity, course), treatment regimen (AD monotherapy or in combination with MSs or APs), pharmacotherapy (AD name, dosage, duration) and outcome-related variables (definition of new depressive episodes (relapse or recurrence) and manic/hypomanic episodes (affective switch), time to new depressive or manic/hypomanic episode, duration of new depressive or manic/hypomanic episodes, dropout rates) were extracted from the included studies. This procedure was performed independently by two investigators (JL, HF) under the guidance of *Cochrane Handbook for Systematic Reviews of Interventions,* any discrepancy detected in the extracted data was settled by discussion among three authors of this paper (BSL, JL, HF). Key missing data or perplex information of included studies were addressed through e-mail contact with the authors.

### 2.5. Risk of bias assessment

Risk of bias of each included trial was assessed according to *Cochrane Handbook for Systematic Reviews of Interventions.* The assessment includes the following six items: randomization generation, allocation concealment, blindness of participants and personnel, blindness of outcome assessment, incomplete outcome data and selective reporting. Risk of bias summary figure was generated according to the above six items by Revman 5.3.

### 2.6. Data synthesis and analysis

Risk ratio (RR) for new depressive episodes was selected as the primary outcome. Risk ratio for new manic/hypomanic episodes, number needed to treat (NNT) and number needed to harm (NNH) derived from risk difference (RD) were selected as secondary outcomes. Due to the exceedingly inconsistent reporting about dropouts in different studies and insufficient data about time to new affective episodes, we finally abandoned meta-analysis of dropout rate, time to new affective episodes and duration of new episodes, leaving the data analysis limited to dichotomous data about new depressive or manic/hypomanic episodes. When different criteria of new depressive, manic/hypomanic episodes were presented in a trial, data about the most formal and commonly used criteria, namely, the criteria of depressive or manic/hypomanic episode based on RDC, DSM or ICD diagnostic system, was selected for meta-analyses. Pooled estimates were tested by Z statistic, and significance was achieved when a two-tailed P value was less than 0.05.

Heterogeneity between included studies was assessed by the Q statistic and the $I^2$ statistics. $P < 0.1$ for Q statistic or $I^2 > 35\%$ was taken as indicator of statistical significant heterogeneity (Higgins and Thompson, 2002). Mantel-Haenszel fixed -effect model (Mantel and Haenszel, 1959) was selected for meta-analyses due to low heterogeneity detected between studies. Subgroup analyses were

implemented to explore whether treatment regimen (monotherapy or combination), bipolar subtypes (I, II or I and II), antidepressant types (first generation or second generation), treatment duration (≤6 m or > 6 m) or funding source (pharmacy or non-pharmacy) exerted influence on the outcomes of efficacy and safety. However, the results of some subgroup analyses were associated with lowered reliability owing to small sample size. Sensitivity and meta-regression were used test the stability of our findings and explore the factors potentially influencing our results. Finally, Funnel plot and Egger's test were applied to explore the potential publication bias.

All the statistic work was performed in the meta-analysis module of Stata 12.0 for Windows.

## 3. Results

### 3.1. Literature search and selection

The electronic database search attached 1508 records and hand search yielded 14 studies. Of these, twenty were retained for full-text screening after duplicates discarded and title and abstract screening. Finally, 9 studies were excluded in full-text screening and eleven studies were included for meta-analyses. The reasons for exclusion in full-text screening were as follows: naturalistic design (Altshuler et al., 2003; Viktorin et al., 2014), study protocol (Simpson et al., 2009), no data reported about the acute responders or remitters (Sachs et al., 2007; Leverich et al., 2006), reanalysis of data from other studies (Amsterdam et al., 2013; Vohringer et al., 2015; Altshuler et al., 2009), inadequate data reporting (Amsterdam et al., 1998). Eleven studies were finally included in the meta-analysis. The study selection procedure was shown in Fig. 1.

### 3.2. Main characteristics of included studies

The comparisons of the included studies were divided into two major groups: comparing AD with placebo (including AD monotherapy and AD combined with MS (AD + MS)) or comparing AD monotherapy with MS monotherapy. The first group (group1) has eleven comparisons from ten trials, including 637 BD patients (322 in AD group and 315 in placebo group). Seven comparisons compared AD + MS with MS + placebo and the other four compared AD alone with placebo. For the simplification of statistical analysis, one study (Brown et al., 2009) comparing Olanzapine-Fluoxetine Combination (OFC) with Lamotrigine was categorized into this group due to the analogous mood-stabilizing effect of Olanzapine and Lamotrigine. The second group (group2) has five comparisons, including 227 BD patients (122 in AD group and 105 in MS group).

The included eleven studies were published from 1973 to 2015, with sample sizes ranging from 13 to 172. The ratio of female of the included subjects ranged from 23% to 77%. Imipramine and fluoxetine were the most commonly studied ADs: five studies (Prien et al., 1984; Prien et al., 1973; Quitkin et al., 1981; Kane et al., 1982; Johnstone et al., 1990) employed imipramine and four (Brown et al., 2009; Tamayo et al., 2009; Amsterdam and Shults, 2010; Amsterdam and Shults, 2005) selected fluoxetine. Only one study hired venlafaxine (Amsterdam et al., 2015) and one study used four ADs (bupropion, paroxetine, citalopram and venlafaxine)(Ghaemi et al., 2010). Lithium Carbonate (Li) is the most frequently used mood stabilizer, except for three studies: one compared OFC with Lamotrigine (Brown et al., 2009), one compared OFC with Olanzapine (Tamayo et al., 2009), and the other has no limitation to a specific MS (Ghaemi et al., 2010). The average completed treatment duration ranged from 4.4 months to 36 months. Note that in two studies (Brown et al., 2009; Amsterdam et al., 2015) the randomization procedure was performed when patients were in acute depressive episode, and data were collected for patients who achieved response and continued their treatment in maintenance phase. Two studies were open trials with one (Ghaemi et al., 2010) compared

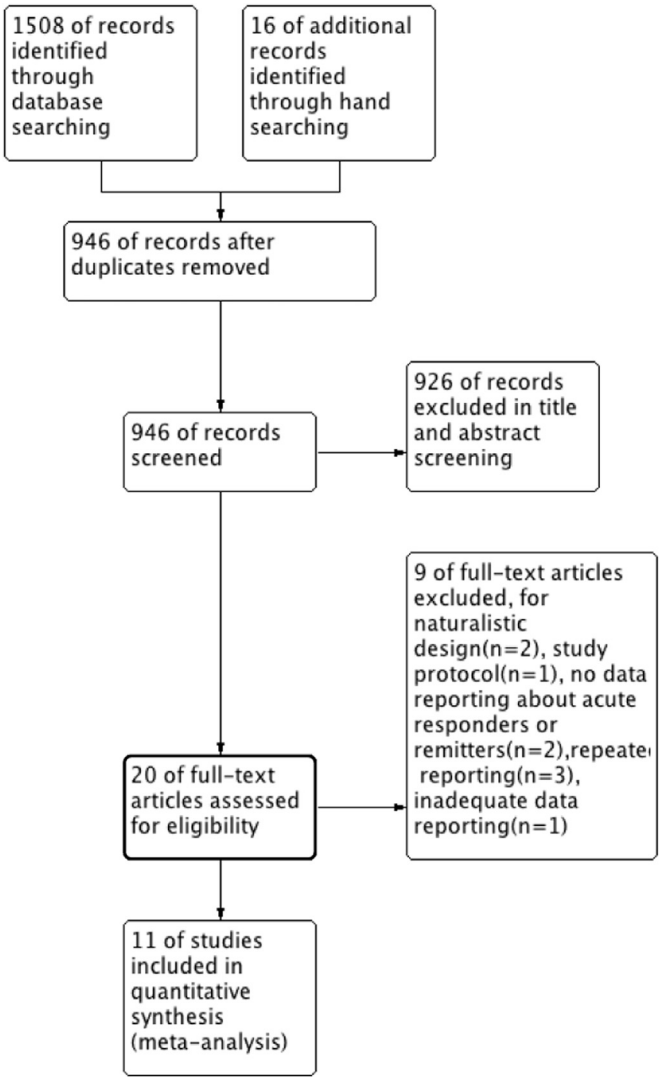


**Fig. 1.** Flowchart of study selection process. Literature was systematically searched from Pubmed, EMBASE, PsyINFO, CENTRAL and clinicaltrials.gov from their first available dates to August 31, 2016. Relevant papers were also handly searched from guidelines, expert consensuses, systematic reviews and meta-analyses.

patients continued with those discontinued AD treatment after remission from acute depression treatment, and the other (Tamayo et al., 2009) compared OFC with Olanzapine (OLZ) in responders of a 7-week OFC treatment for acute depressive episode. All studies except one (Johnstone et al., 1990) recruited enriched sample, which means only patients who achieved benefit from acute AD or MS therapy were included in the long-term continuation or maintenance treatment. The main characteristics of the included studies are demonstrated in Table 1.

### 3.3. Methodological quality

All the eleven included studies reported randomized controlled trial design, while only two studies (Amsterdam et al., 2015; Ghaemi et al., 2010) explicitly declared the method of random sequence generating. No study reported the specific method of allocation concealment. Three studies were evaluated as high risk of bias, owing to non-blindness of the patients, treating clinicians and/or evaluators (Johnstone et al., 1990; Tamayo et al., 2009; Ghaemi et al., 2010) and incomplete outcome data reporting (Ghaemi et al., 2010)(Appendix). There is hardly any bias of selective reporting of the included eleven trials.

**Table 1**
Main characteristics of included studies.

| Study | Funding | N | F% | Age (mean ± SD) | Index episode | Inclusion criteria | Treatments (dose) | Duration (m) | Enriched sample? | New depressive episode criteria | New manic or hypomanic episode criteria | Included comparisons | Dropout rate (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Prien 1973 | Fed | 44 | 23 | 43.3 ± 11.9 | D | Hosp for acute MDE; ≥1 AE for hosp in the past 2 years | Imi (125 [50–200]); Li (1250 [500–2250];0.5–1.4); Pbo | 24(-) | Y | depressive attack needs hosp or extra drugs | manic attack needs hosp or extra drugs | Imi vs placebo; Imi vs Li | 54.5 |
| Quitkin 1981 | NS | 75 | 52 | 36.8 ± 13.1 | NS | RDC BD I; euthymic for 6w while receiving only Li | Li + Imi (100–150) Pbo + Li (0.8–1.2) | 24 (18.8) | Y | RDC MDE for 1w or RDC minor MDE for 4w | RDC mania or RDC hypomania for 1w | Imi + Li vs Li + placebo | 5.3 |
| Kane 1982 | Fnd | 22 | 77 | 48.7(-) | D | RDC BD II; ≥2 AEs in the past 7 years; euthymic for 6 m | Imi + Li (0.8–1.2) Pbo + Imi (100–150) Pbo + Imi(100–150) Pbo + Pbo | 19 (11) | Y | RDC MDE for 1w or RDC minor MDE for 4w | RDC mania or RDC hypomania for 1w | Imi + Li vs Li + placebo; Imi vs placebo; Imi vs Li | 45.5* |
| Prien 1984 | Fed | 114 | 58 | 38.1 ± 12.4 | D or M | RDC manic disorder; ≥7 in RSDM; ≤60 in GAS; ≥1 AE in the past 2.5 years | Li + Imi (132 [75–150]) Li (0.75 [0.45– 1.1]) | 24 (-) | Y | RDC MDE; ≤60 in GAS | RDC mania; ≤60 in GAS | Imi + Li vs Li + placebo; Imi vs Li | 0.9 |
| Johnstone 1990 | Pharm | 13 | 69 | 50.5 ± 9.5 | D | DSM-III BD | Li + AMI (88 [50–124]) Li (0.5–0.9) | 36 (9.6) | N | DSM-III-R MDE | DSM-III-R mania | Imi + Li vs Li + placebo | 46.2 |
| Amsterdam 2005 | Fnd + Pharm | 12 | 75 | 38.0 ± 14.2 | D | DSM-IV MDE; BD II or NOS; ≥18 in HRSD − 17 | Fluo (20); Pbo | 8 (-) | Y | ≥14 increase in HRSD − 17 and DSM-IV MDE | / | Fluo vs placebo | 0.0 |
| Brown 2009 | Pharm | 172 | 60 | 36.8 ± 11.5 | D | DSM-IV BD I MDE; ≥20 on MADRS; ≥4 on GGI-S; ≥1 AE requiring treatment with MSs or APs | OFC (Olz/ Fluo: 10.7/ 40.8) OLZ (10.7) | 5.8 | Y | MADRS≥15 | / | OFC vs Lam | NA** |
| Tamayo 2009 | Pharm | 114 | 66 | 44.5 ± 11.3 | D | DSM-IV-TR BD I or II; MADRS≥20 | OFC (Olz/ Fluo: 10.3/ 33.6) Olz (10.8) | 4.4 | Y | MADRS≥20; CGI-BP-D≥3 | CGI-BP mania≥13 or hosp for mania | OFC vs OLZ | 27 |
| Ghaemi 2010 | Fed | 70 | 53 | 43.2 ± 13.2 | D | DSM-IV BD I, II, or NOS; 70% BD I | ADs + MS MS (std Rxs) | 36(-) | Y | DSM-IV MDE | DSM-IV mania | AD + MS vs MS alone | 61 |
| Amsterdam 2010 | Fed + Fnd | 81 | 52 | 37.7 ± 12.1 | D | DSM-IV-TR BD II MDE; ≥16 in HRSD − 17 | Fluo (34.3 ± 7.9) Li (1027 ± 210.8; 0.69 ± 0.27) Pbo | 15.5 | Y | ≥14 in HRSD and DSM-IV MDE | DSM-IV hypomania | Fluo vs placebo | 22.2 |
| Amsterdam 2015 | Fed + Fnd | 55 | 54 | 42.2.7 ± 12.5 | D | DSM-IV-TR BD II MDE; ≥16 in HRSD − 17 | Venl (253.9 [112.5–375]) Li (1090 [300–1740]; 0.72 [0.30–1.08]) | 8.8 | Y | DSM-IV MDE | DSM-IV hypomania | Venl vs Lithium | 18.2 |

Fed: Federal; Pharma: Pharmacy; Fnd: funding of foundations; NS: not stated; D: depressive; M: manic/hypomanic; BD: bipolar disorder; AE: affective episode; MS: Mood Stabilizer; AP: Antipsychotic; RDC: Research Domain Criteria; DSM: Diagnostic and Statistical Manual of mental disorders; MDE: Major Depressive Episode; HRSD: Hamilton Rating Scale for Depression; MADRS: Montgomery-Asberg Depression Rating Scale; CGI-BP: Clinical Global Impression scale for Bipolar Disorder; NS: not sated; hosp: hospitalization; Imi: Imipramine; Li: Lithium Carbonate; AD: antidepressant; OFC: Olanzapine-Fluoxetine Combination; OLZ: Olanzapine; Fluo: Fluoxetine; Venl: Venlafaxine. The doses of the drugs were presented in range or mean ± SD, and the both the doses and serum concentration of Li was presented.

* Proportion of patients terminated with euthymic state.

** The dropouts for the patients enrolled in the study between the two groups were similar: 66.8% for OFC group and 66.3% for Lam group. While the dropouts for the patients achieved remission were not available.

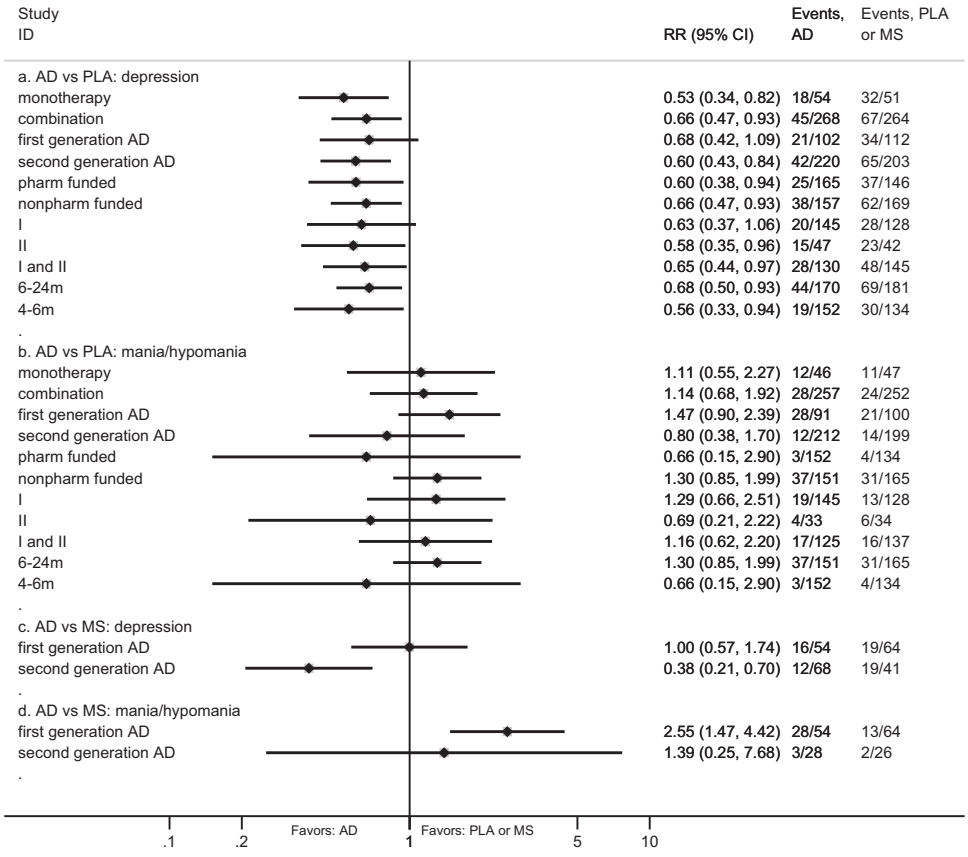


**Fig. 2.** Meta-analyses of AD vs PLA or MS in prophylaxis of new depressive episodes or inducement of new manic/hypomanic episodes. a. AD vs PLA in prophylaxis of new depressive episodes; b. AD vs PLA in risk of new manic/hypomanic episodes; c. AD vs MS in prophylaxis of new depressive episodes; d. AD vs MS in risk of new manic/hypomanic episodes. Abbreviations: AD: antidepressants; PLA: placebo; MS: mood stabilizers; RR: Risk Ratio; CI: confidence interval.

### *3.4. AD vs placebo (including AD + MS and AD monotherapy)*

#### *3.4.1. Prophylaxis of new depressive episodes*

Ten studies with 637 BD patients reported data comparing the efficacy of AD with placebo in prophylaxis of new depressive episodes, with seven comparisons (N = 532) compared AD in combination with MS (AD + MS) with MS (+ placebo) and four comparisons (N = 105) compared AD monotherapy with placebo. The relapse rate of new depressive episodes was significantly lower in AD group than in placebo group (RR = 0.64, 95% confidence interval (CI) [0.49–0.83], P = 0.0009) (Fig. 2a), either AD + MS (RR = 0.70, 95%CI [0.50–0.97], P = 0.03) or AD monotherapy (RR = 0.53, 95%CI [0.35–0.81], P = 0.003) subgroup reached statistical significance (Fig. 3**a**). The AD monotherapy regimen revealed a relatively higher efficacy than AD + MS, though not statistically significant (P = 0.31). The NNT for the total synthesis was 9.1, 95% CI [5.6–20] (for AD + MS: NNT = 12.5, 95% CI [7.1–100]; for AD monotherapy: NNT = 3.3 95% CI [2.0–8.3]), indicating overall moderate efficacy of AD for prevention of new depressive episodes, with high efficacy in AD monotherapy and relatively small efficacy in AD + MS regimen.

The heterogeneity test showed little heterogeneity between the studies (Q = 6.75, df = 10, P = 0.749, $I^2$ = 0%). By removing one trial each time, the meta-influence analysis didn't reveal any outlier tended to bias the results substantially (Appendix). The funnel plot is roughly symmetric and the Egger's intercept covers zero (t = 0.16, p = 0.880, 95% CI [−1.75, 2.01]), indicating no signs of publication bias (Appendix).

#### *3.4.2. Risk of inducement of new manic/hypomanic episodes*

As no patients switched to mania or hypomania in either AD or placebo group in three studies (Kane et al., 1982; Johnstone et al., 1990; Amsterdam and Shults, 2005), these studies were not included in the meta-analysis of new manic/hypomanic episodes. Finally, data from eight comparisons with 602 BD patients (303 in AD group and 299 in placebo group) were synthesized. The pooled RR was 1.21, 95% CI [0.82–1.80] (for AD + MS: RR = 1.26, 95% CI [0.77–2.05]; for AD monotherapy: RR = 1.11, 95% CI [0.57–2.16]) (Fig. 2b; 3b). The RD for the total synthesis was 0.02, 95%CI [−0.03–0.07], P = 0.35, NNH = 50 (for AD + MS: RD = 0.02, 95% CI [−0.03–0.08]), P = 0.35, NNH = 50; for AD monotherapy: RD = 0.03, 95% CI = −0.14–0.19, P = 0.76, NNH = 33), indicating no significant difference between AD and placebo (either AD + MS or AD monotherapy) in inducing new manic/hypomanic episodes. No significant heterogeneity was detected between the included studies (Q = 0.18, df = 7, P = 0.638, $I^2$ = 0%).

### *3.5. AD monotherapy vs MS monotherapy*

#### *3.5.1. Prophylaxis of new depressive episodes*

Five trials (Prien et al., 1984; Prien et al., 1973; Kane et al., 1982; Amsterdam and Shults, 2010; Amsterdam et al., 2015) with 227 patients (122 in AD group and 105 in MS group) compared AD monotherapy with MS monotherapy in prophylaxis of new depressive episodes and inducement of new manic/hypomanic episodes. All the five trials were included in the meta-analysis of new depressive episodes. No significant difference between AD monotherapy and MS monotherapy in prevention of new depressive episodes was detected (RR = 0.71, 95%CI [0.48–1.06], P = 0.09; RD = −0.10, 95% CI [−0.22-0.02]) (Fig. 2c). Heterogeneity between these studies was low (Q = 3.93, df = 4, P = 0.415, $I^2$ = 0%).

#### *3.5.2. Risk of inducement of new manic/hypomanic episodes*

As no patient experienced new mania/hypomania episodes in either AD or placebo group in the study Amsterdam et al., 2015 (Amsterdam et al., 2015), this study was not included in the analysis of new manic/hypomanic episodes. The pooled estimate of the remained four studies (N = 172) revealed significantly higher risk of new manic/hypomanic episodes in AD monotherapy than MS monotherapy (RR = 2.35, 95%CI [1.42–3.91], P = 0.001) (Fig. 2d). The RD between two groups is large

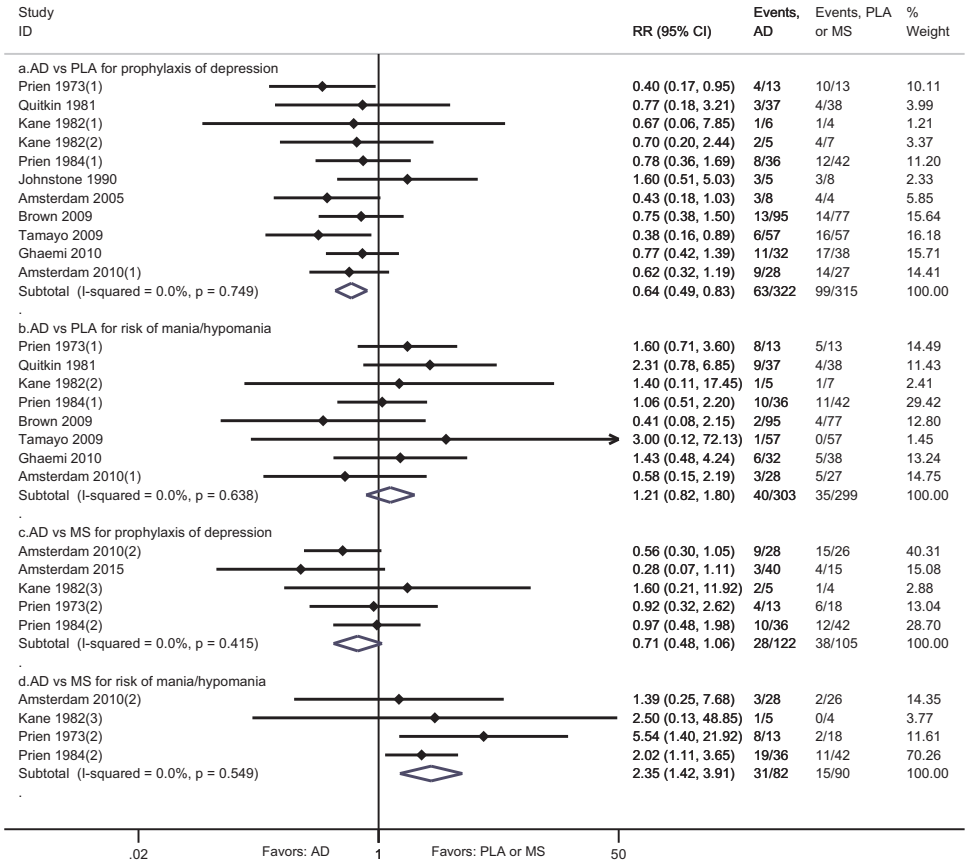


**Fig. 3.** Summary of subgroup meta-analyses comparing AD with PLA or MS in prophylaxis of new depressive episodes or inducement of new manic/hypomanic episodes. a. AD vs PLA in prophylaxis of new depressive episodes; b. AD vs PLA in risk of new manic/hypomanic episodes; c. AD vs MS in prophylaxis of new depressive episodes; d. AD vs MS in risk of new manic/hypomanic episodes. These meta-analyses were subgrouped by treatment regimen (monotherapy vs combined with MS), antidepressant types (first generation vs second generation), funding source (pharmacy funded vs non-pharmacy funded), BD subtypes (I vs II vs I and II) and treatment duration (6–24 m vs 4–6 m). Abbreviations: AD: antidepressants; PLA: placebo; MS: mood stabilizers; RR: Risk Ratio; CI: confidence interval.

(RD = 0.23,95% CI [0.11–0.35], P = 0.0002), and the NNH is small (NNH = 4.3, 95% CI [2.9–9.1]), indicating significantly higher risk of affective switch in AD monotherapy comparing with MS monotherapy. The heterogeneity in this analysis is also low (Q = 2.11, df = 3, P = 0.549, $I^2$ = 0%).

### 3.6. Subgroup analyses and meta-regression

Subgroup analyses were done by BD subtypes (I, II, I and II), AD types (first generation, second generation), funding source (pharmacy, non-pharmacy) and treatment duration (≤6 m, > 6 m). These subgroups revealed that only BD II benefits more from long-term AD treatment: BD II and BD I + II subgroups had significantly reduced risk of new depressive episodes, while BD I subgroup had only slightly and not significantly diminished risk of new depression (Fig. 3a); BD I subgroup exhibited moderately higher RR of new manic/hypomanic episodes than BD II, though this difference didn't achieve statistical significance (Fig. 3b). Only second generation AD exhibited significant reduction in new depressive episodes in BD patients, while first generation AD did not (both first and second generation AD subgroup didn't increase the risk of new manic/hypomanic episodes)( (Fig. 3a & b), this may be due to the fact that most trials used first generation ADs also recruited BD I or BD I and II rather than BD II). The relatively small sample size (N = 214) in this subgroup may also contribute to nonsignificance. The funding source and treatment duration exhibited no significant influence on the efficacy of depression prophylaxis and mania/hypomania inducement (Fig. 3a & b). Meta-regression analysis of sample size and sex ratio didn't reveal statistical significance (Appendix).

## 4. Discussion

Long-term antidepressant use in BD patients is commonly seen in clinical practice, while evidence for this phenomenon is scant (Grunze et al., 2013). In this study, we meta-analyzed the available published RCTs addressing the efficacy and safety of long-term (≥4 m) AD treatment for BD patients. Our meta-analyses found that: (1) as compared with placebo, long-term AD treatment, either used as monotherapy or in combination with MS, is effective for prophylaxis of new depressive episodes without increasing the risk of new manic/hypomanic episodes; (2) as compared with MS monotherapy, long-term AD monotherapy didn't reveal any superiority in preventing new depressive episodes, but had significantly increased risk of new mania/hypomania; (3) BD II subtype benefits more from long-term AD treatment than BD I.

Psychiatrists were often puzzled by the benefit and risk of prescribing AD for preventing new depressive episodes in bipolar patients. Our meta-analyses supported the benefit of long-term AD treatment in reducing new depressive episodes in bipolar disorder patients, either used alone or in combination with MS. However, smaller efficacy of AD + MS regimen (NNT = 12.5) as compared with AD monotherapy (NNT = 3.3) was detected, indicating potential synergetic effect of AD and MS in preventing depression relapse and recurrence and minor superiority of MS + AD regimen over MS monotherapy.

Beyond disputes about efficacy, the risk of long-term AD treatment in bipolar disorder is often more contentious. The findings of no increased risk of mania/hypomania during long-term AD treatment as compared with placebo while elevated risk of affective switch as compared with MS monotherapy were interesting. The use, especially long-term use, of AD in BD patients, is intensively augmentative due to the potential risk of mania/hypomania inducement. However, our results indicated that the affective switch of BD may be mainly contributed to the naturalistic course of the disease itself rather than AD's inducement, and, the lower risk of new manic/hypomanic episodes in MS monotherapy group as compared with AD monotherapy group may be explained by the protective effect of MS in preventing new mania/

hypomania symptoms.

The previous meta-analysis (Ghaemi et al., 2008) addressing the benefit and risk of long-term AD treatment for BD found little benefit in preventing new depressive episodes while great risk in inducing new manic/hypomanic episodes, which is contradicted with our findings. This discrepancy may be interpreted by the separation of comparisons between AD and placebo and comparisons between AD and MS in our study. As a matter of fact, our findings are consistent with the results of subgroup analyses of AD only vs placebo only and AD only vs MS only in the previous meta-analysis, which also revealed no difference in risk of new manic/hypomanic episodes between AD (+MS) and placebo (+MS) while significantly elevated risk in AD only as compared with MS only.

The findings about the efficacy and safety of long-term AD+MS treatment for BD is also congruent with conclusions of two previous naturalistic observational studies (Altshuler et al., 2003; Viktorin et al., 2014), which declared that benefits of long-term AD+MS treatment in diminishing risk of new depressive episodes can be achieved with no increase in risk of new manic/hypomanic episodes. However, the study of Viktorin et al.Viktorin et al. (2014) also found increased risk of affective switch in patients receiving AD monotherapy, which is opposite to our findings. A possible explanation for this discrepancy may lie in the fact that most comparisons comparing AD monotherapy with placebo included in our study recruited BD II patients (three (Kane et al., 1982; Amsterdam and Shults, 2010; Amsterdam and Shults, 2005) of four recruited only BD II patients, the other (Prien et al., 1973) didn't declare details of BD subtypes), while the study of Viktorin et al. included both patients with BD I and BD II. Since BD II is characterized by predominantly depressive episodes, the risk of mania/hypomania inducement in this subtype may be much less than that in BD I (Vohringer et al., 2015; Amsterdam et al., 2015).

Considering the different clinical manifestation between BD I and BD II, whether clinicians should adopt distinct strategies in prescribing AD for the two subtypes still remains disputable. As illustrated in the subgroup meta-analyses, BD II subgroup demonstrated moderately higher RR in new manic/hypomanic episodes and slightly lower RR in new depressive episodes than BD I subgroup, although both indicators didn't achieve statistical significance in test for subgroup differences, the trend of greater benefit and lower risk of long-term AD treatment in BD II than in BD I may call for attention in differentiating treatment strategies for the two subtypes.

Despite the exhilarating findings of our study, some limitations call for caution in interpreting the results: (1) The number of included studies is only eleven and the total sample size synthesized is only moderate (N=692); (2) Although all the eleven included studies were RCTs, the randomization procedure was performed in acute depression phase rather than patients who achieved remission in two studies (Brown et al., 2009; Amsterdam et al., 2015), which may introduce selection bias; (3) Three studies included patients (Tamayo et al., 2009; Amsterdam et al., 2015; Ghaemi et al., 2010) achieved only response rather than remission in acute depression phase, and the treatment durations were less than six months in two studies (Brown et al., 2009; Tamayo et al., 2009), which do not meet the standard criteria of prophylaxis (maintenance) treatment of BD(Tohen et al., 2009; American Psychiatric Association, 2013) and may confound the residual symptoms of one episode with another new episode; (4) The non-blindness of patients, treating clinicians and/or evaluators in three included studies (Johnstone et al., 1990; Tamayo et al., 2009; Ghaemi et al., 2010) may introduce performance and measurement bias, and high dropout rate in one studies (Ghaemi et al., 2010) is associated with attrition bias; (5) All studies except one (Johnstone et al., 1990) selected enriched sample, which may limit the external validity of the results of these meta-analyses; (6) Owing to inadequate or inconsistent data reporting, we finally abandoned the planned meta-analyses of time to relapse, duration of relapse and dropouts, which are also of key value for us to understand the efficacy and safety of long-term AD treatment.

Considering the above limitations, future RCTs addressing the risk and benefit of long-term AD treatment for bipolar disorder should implement randomization in patients who had achieved remission of depression and report clearly about time to depression relapse/recurrence or mania/hypomania, duration of new affective episodes and dropouts. And, blindness of participants, treating clinicians and data analyzers should be tried best to achieve. What's more, more studies are needed to address the risk and benefits of long-term AD treatment in non-enriched BD samples.

## 5. Conclusions

Our meta-analyses suggested that the evidence from current RCTs seems to back up the long-term use of AD in BD treatment, owing to its efficacy in prophylaxis of new depressive episodes without significantly increased risk of mania/hypomania inducement. The higher risk of new manic/hypomanic episodes in AD monotherapy as compared with MS monotherapy may be derived from the protective efficacy of MS for preventing new mania/hypomanic episodes. The superiority of AD+MS over MS alone, although relatively small, may justify the commonly seen phenomenon of long-term AD treatment in BD patients in clinical practice. Longer AD treatment than recommended in guidelines should be considered for some BD patients, especially those with BD II disorder. On account of the limitations of our meta-analyses, future studies are imperatively needed to address this issue further on.

## Conflict of interest

The authors declare no conflict of interest.

## Author disclosure

None.

## Contributors

LJL and TBL co-designed the topic. LJL organized team work and had complete access to all data of this paper. BSL and YZ discussed and determined the search strategy. HF and JL performed literature screening and data extraction separately, their disagreements were discussed and solved with BSL. YZ and HF did the meta-analyses and produced the figures. BSL drafted the manuscript and YZ supplemented modifications and polishments. All the authors have reviewed and approved the final version of the manuscript.

## Role of the funding source

The authors acknowledge financial support from the National Science and Technologic Program of China (2015BAI13B02 to Lingjiang Li), National Basic Research Program of China (+2013CB835100 to Lingjiang Li) and National Natural Science Foundation of China (81171286 & 91232714 to Lingjiang Li; 81671353 to Yan Zhang). The sponsors had no role in the study design, data collection and analysis, interpretation of the results and writing of the report.

## Acknowledgements

None.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jad.2017.07.023.

## References


Bopp, J.M., Miklowitz, D.J., Goodwin, G.M., et al., 2010. The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study. Bipolar Disord. 12 (3), 327–334.

Judd, L.L., Akiskal, H.S., Schettler, P.J., et al., 2002. The long-term natural history ofthe weekly symptomatic status of bipolar I disorder. Arch. Gen. Psychiatry 59 (6), 530–537.

Miller, S., Dell'Osso, B., Ketter, T.A., 2014. The prevalence and burden of bipolar depression. J. Affect Disord. 169 (Suppl 1S3–11).

Michalak, E.E., Murray, G., Young, A.H., et al., 2008. Burden of bipolar depression: impact of disorder and medications on quality of life. CNS Drugs 22 (5), 389–406.

Solomon, D.A., Leon, A.C., Coryell, W.H., et al., 2010. Longitudinal course of bipolar I disorder: duration of mood episodes. Arch. Gen. Psychiatry 67 (4), 339–347.

Grunze, H., Vieta, E., Goodwin, G.M., et al., 2013. The World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for the biological treatment of bipolar disorders: update 2012 on the long-term treatment of bipolar disorder. World J. Biol. Psychiatry 14 (3), 154–219.

Goodwin, G.M., 2009. Consensus Group of the British Association for P. Evidence-based guidelines for treating bipolar disorder: revised second edition–recommendations from the British Association for Psychopharmacology. J. Psychopharmacol. 23 (4), 346–388.

Yatham, L.N., Kennedy, S.H., Parikh, S.V., et al., 2013. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013[J]. Bipolar Disord. 15 (1), 1–44.

Malhi, G.S., Bassett, D., Boyce, P., et al., 2015. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Aust. N. Z. J. Psychiatry 49 (12), 1087–1206.

Pacchiarotti, I., Bond, D.J., Baldessarini, R.J., et al., 2013. The International Society for Bipolar Disorders (ISBD) task force report on antidepressant use in bipolar disorders. Am. J. Psychiatry 170 (11), 1249–1262.

Davidson, J.R., 2010. Major depressive disorder treatment guidelines in America and Europe. J. Clin. Psychiatry 71 (Suppl) (E1e04).

Grande, I., de Arce, R., Jimenez-Arriero, M.A., et al., 2013. Patterns of pharmacological maintenance treatment in a community mental health services bipolar disorder cohort study (SIN-DEPRES). Int J. Neuropsychopharmacol. 16 (3), 513–523.

Gitlin, M.J., 2012. Antidepressants in bipolar depression: much confusion, many questions, few answers. Aust. N. Z. J. Psychiatry 46 (4), 295–297.

Ghaemi, S.N., 2012. Antidepressants in bipolar depression: the clinical debate. Aust. N. Z. J. Psychiatry 46 (4), 298–301.

Ghaemi, S.N., Wingo, A.P., Filkowski, M.A., et al., 2008. Long-term antidepressant treatment in bipolar disorder: meta-analyses of benefits and risks. Acta Psychiatr. Scand. 118 (5), 347–356.

Sidor, M.M., Macqueen, G.M., 2011. Antidepressants for the acute treatment of bipolar depression: a systematic review and meta-analysis. J. Clin. Psychiatry 72 (2), 156–167.

Higgins, J.P., Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. Stat. Med 21 (11), 1539–1558.

Mantel, N., Haenszel, W., 1959. Statistical aspects of the analysis of data from retrospective studies of disease. J. Natl. Cancer Inst. 22 (4), 719–748.

Altshuler, L., Suppes, T., Black, D., et al., 2003. Impact of antidepressant discontinuation after acute bipolar depression remission on rates of depressive relapse at 1-year follow-up. Am. J. Psychiatry 160 (7), 1252–1262.

Viktorin, A., Lichtenstein, P., Thase, M.E., et al., 2014. The risk of switch to mania in patients with bipolar disorder during treatment with an antidepressant alone and in combination with a mood stabilizer. Am. J. Psychiatry 171 (10), 1067–1073.

Simpson, S., Barnes, E., Griffiths, E., et al., 2009. The Bipolar Interactive Psychoeducation (BIPED) study: trial design and protocol. BMC Psychiatry 950.

Sachs, G.S., Nierenberg, A.A., Calabrese, J.R., et al., 2007. Effectiveness of adjunctive antidepressant treatment for bipolar depression. N. Engl. J. Med 356 (17), 1711–1722.

Leverich, G.S., Altshuler, L.L., Frye, M.A., et al., 2006. Risk of switch in mood polarity to hypomania or mania in patients with bipolar depression during acute and continuation trials of venlafaxine, sertraline, and bupropion as adjuncts to mood stabilizers. Am. J. Psychiatry 163 (2), 232–239.

Amsterdam, J.D., Luo, L., Shults, J., 2013. Efficacy and mood conversion rate during long-term fluoxetine v. lithium monotherapy in rapid- and non-rapid-cycling bipolar II disorder. Br. J. Psychiatry 202 (4), 301–306.

Vohringer, P.A., Ostacher, M.J., El-Mallakh, R.S., et al., 2015. Antidepressants in Type II Versus Type I Bipolar Depression: A Randomized Discontinuation Trial. J. Clin. Psychopharmacol. 35 (5), 605–608.

Altshuler, L.L., Post, R.M., Hellemann, G., et al., 2009. Impact of antidepressant continuation after acute positive or partial treatment response for bipolar depression: a blinded, randomized study. J. Clin. Psychiatry 70 (4), 450–457.

Amsterdam, J.D., Garcia-Espana, F., Fawcett, J., et al., 1998. Efficacy and safety of fluoxetine in treating bipolar II major depressive episode. J. Clin. Psychopharmacol. 18 (6), 435–440.

Brown, E., Dunner, D.L., McElroy, S.L., et al., 2009. Olanzapine/fluoxetine combination vs. lamotrigine in the 6-month treatment of bipolar I depression. Int J. Neuropsychopharmacol. 12 (6), 773–782.

Prien, R.F., Kupfer, D.J., Mansky, P.A., et al., 1984. Drug therapy in the prevention of recurrences in unipolar and bipolar affective disorders. Report of the NIMH Collaborative Study Group comparing lithium carbonate, imipramine, and a lithium carbonate-imipramine combination. Arch. Gen. Psychiatry 41 (11), 1096–1104.

Prien, R.F., Klett, C.J., Caffey Jr., E.M., 1973. Lithium carbonate and imipramine in prevention of affective episodes. A comparison in recurrent affective illness. Arch. Gen. Psychiatry 29 (3), 420–425.

Quitkin, F.M., Kane, J., Rifkin, A., et al., 1981. Prophylactic lithium carbonate with and without imipramine for bipolar 1 patients. A double-blind study. Arch. Gen. Psychiatry 38 (8), 902–907.

Kane, J.M., Quitkin, F.M., Rifkin, A., et al., 1982. Lithium carbonate and imipramine in the prophylaxis of unipolar and bipolar II illness: a prospective, placebo-controlled comparison. Arch. Gen. Psychiatry 39 (9), 1065–1069.

Johnstone, E.C., Owens, D.G., Lambert, M.T., et al., 1990. Combination tricyclic antidepressant and lithium maintenance medication in unipolar and bipolar depressed patients. J. Affect Disord. 20 (4), 225–233.

Tamayo, J.M., Sutton, V.K., Mattei, M.A., et al., 2009. Effectiveness and safety of the combination of fluoxetine and olanzapine in outpatients with bipolar depression: an open-label, randomized, flexible-dose study in Puerto Rico. J. Clin. Psychopharmacol. 29 (4), 358–361.

Amsterdam, J.D., Shults, J., 2010. Efficacy and safety of long-term fluoxetine versus lithium monotherapy of bipolar II disorder: a randomized, double-blind, placebo-substitution study. Am. J. Psychiatry 167 (7), 792–800.

Amsterdam, J.D., Shults, J., 2005. Fluoxetine monotherapy of bipolar type II and bipolar NOS major depression: a double-blind, placebo-substitution, continuation study. Int Clin. Psychopharmacol. 20 (5), 257–264.

Amsterdam, J.D., Lorenzo-Luaces, L., Soeller, I., et al., 2015. Safety and effectiveness of continuation antidepressant versus mood stabilizer monotherapy for relapse-prevention of bipolar II depression: a randomized, double-blind, parallel-group, prospective study. J. Affect Disord. 18531–18537.

Ghaemi, S.N., Ostacher, M.M., El-Mallakh, R.S., et al., 2010. Antidepressant discontinuation in bipolar depression: a Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) randomized clinical trial of long-term effectiveness and safety. J. Clin. Psychiatry 71 (4), 372–380.

Tohen, M., Frank, E., Bowden, C.L., et al., 2009. The International Society for Bipolar Disorders (ISBD) Task Force report on the nomenclature of course and outcome in bipolar disorders. Bipolar Disord. 11 (5), 453–473.

American Psychiatric Association, Diagnostic and statistical manual of mental disorders. 5th edition 2013, American Psychiatric Press Washington, DC.