Psychological Medicine

cambridge.org/psm

Original Article

‡The online version of this article has been updated since original publication. A notice detailing the changes has also been published at https://doi.org/10.1017/S0033291721001033

Cite this article: Kishi T, Matsuda Y, Sakuma K, Okuya M, Mishima K, Iwata N (2021). Recurrence rates in stable bipolar disorder patients after drug discontinuation *v.* drug maintenance: a systematic review and meta-analysis. *Psychological Medicine* **51**, 2721–2729. https://doi.org/10.1017/S0033291720003505

Received: 3 August 2020
Revised: 26 August 2020
Accepted: 7 September 2020
First published online: 13 October 2020

Key words:
Antipsychotics; mood stabilizer; recurrence rate; systematic review and meta-analysis; double-blind, randomized placebo-controlled trials

Author for correspondence:
Taro Kishi, E-mail: tarok@fujita-hu.ac.jp

CAMBRIDGE UNIVERSITY PRESS

Recurrence rates in stable bipolar disorder patients after drug discontinuation *v.* drug maintenance: a systematic review and meta-analysis‡

Taro Kishi[1], Yuki Matsuda[2], Kenji Sakuma[1], Makoto Okuya[1], Kazuo Mishima[3] and Nakao Iwata[1]

[1]Department of Psychiatry, Fujita Health University School of Medicine, Toyoake, Aichi 470-1192, Japan; [2]Department of Psychiatry, Jikei University School of Medicine, Minato-ku, Tokyo, 105-8461, Japan and [3]Department of Neuropsychiatry, Akita University Graduate School of Medicine, Akita, Japan

Abstract

Background. This random-effects model meta-analysis of double-blind, randomized placebo-controlled trials compared recurrence rates in bipolar disorder (BD) patients between antipsychotic/mood stabilizer discontinuation and maintenance groups.

Methods. We conducted systematic literature search of Embase, PubMed, and CENTRAL databases without language restriction from inception until 22 May 2020. Independent investigators assessed studies and extracted data. We calculated risk ratios (RRs) and numbers needed to benefit or harm (NNTB/NNTH). Primary outcome was the recurrence rate of any mood episode at 6 months. Secondary outcomes were recurrence rates of depressive episodes and manic/hypomanic/mixed episodes and all-cause discontinuation at 6 months. We also investigated these outcomes at 1, 3, 9, 12, 18, and 24 months.

Results. We identified 22 studies ($n = 5462$) receiving aripiprazole, asenapine, divalproex, long-acting injectable (LAI)-aripiprazole, LAI-risperidone, lamotrigine, lithium, olanzapine, paliperidone, or quetiapine. Mean study duration was 64.50 ± 69.35 weeks. The maintenance group demonstrated lower recurrence rates of any mood episode, depressive episodes, and manic/hypomanic/mixed episodes as well as reduced all-cause discontinuation at every observational point. The RRs (95% confidence interval, NNTB/NNTH) of recurrence rate at 6 months were 0.61 (0.54–0.70, 5) for any mood episode, 0.72 (0.60–0.87, 13) for depressive episodes, and 0.45 (0.36–0.57, 6) for manic/hypomanic/mixed episodes. The RR for all-cause discontinuation at 6 months was 0.71 (0.61–0.82, 6).

Conclusions. Maintaining drug treatment during clinically stable BD prevented recurrence for up to 24 months. Discontinuation of medications for ⩾1 month significantly increased recurrence risk. However, 47.3% of patients who discontinued drugs for 6 months did not experience recurrence.

Introduction

Bipolar disorder (BD) is a common chronic mental disorder with a worldwide prevalence of ≥1%. Patients suffer from repeated but irregular manic/hypomanic episodes, mixed affective state, or depressive episodes throughout their life (Grande, Berk, Birmaher, & Vieta, 2016). Thus, they should continue long-term treatment (especially pharmacological treatment), in order to prevent relapse or recurrence and reduce symptoms (Grande et al., 2016). Recent BD treatment guideline recommended various second generation antipsychotics (SGAs) and mood stabilizers (MSs) as first-line treatment of BD in the maintenance (stable) phase (Yatham et al., 2018), but clinicians may discontinue treatment for various reasons such as intolerable side effects. However, it remains unclear how soon a clinically significant increase in relapse or recurrence risk emerges among patients with clinically stable BD following discontinuation compared with medication maintenance. Therefore, we conducted a systematic review and pair-wise meta-analysis of double-blind, randomized placebo-controlled trials (DBRPCTs) comparing recurrence rates between medication discontinuation and maintenance groups of BD patients at 1, 3, 6, 9, 12, 18, and 24 months. We also conducted a single-group summary meta-analysis to calculate the exact recurrence rates at these time points, both in the maintenance and discontinuation groups.

Long-acting injectable (LAI) antipsychotics possess several benefits over oral antipsychotics. These include more stable blood levels, consistent bioavailability, predictable medication adherence, and an improved pharmacokinetic profile, all of which allow for lower dosages (Kishi, Oya, & Iwata, 2016). Therefore, we also conducted a subgroup analysis to assess

whether the difference in the assigned drug (e.g. LAI-SGA *v.* oral medication) influenced the effect size for recurrence rate and the exact recurrence rates in the maintenance group. Moreover, we performed additional subgroup analysis to investigate whether differences in the characteristics of medications used before randomization (such as drug half-life, online Supplementary Table S1) influenced the exact event rate of recurrence in the discontinuation group.

Methods

We performed this systematic review and network meta-analysis according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (Moher, Liberati, Tetzlaff, Altman, & Group, 2009). The literature search, data extraction, and entry into a spreadsheet for analysis were simultaneously and independently conducted by at least two authors (among TK, KS, MO, and YM). The authors double checked all data for accuracy. Any discrepancies between authors were resolved by discussion with a third author (KM and NI). The study was registered with Open Science Framework (https://osf.io/zxsc5).

Search strategy and inclusion criteria

The literature search and selection flow is illustrated in online Supplementary Fig. S1. We performed a systematic literature review according to the PICO strategy (Patients: adult patients with BD in the maintenance phase, Intervention: monotherapy of antipsychotics and/or monotherapy of MSs, Control: placebo, Outcomes: see the following section). Inclusion criteria were (1) DBRPCTs of antipsychotics and/or MSs lasting at least 12 weeks, (2) DBRPCTs including adult patients with any BD subtype in the maintenance phase, (3) DBRPCTs including patients with any mood symptoms at recruitment, and (4) DBRPCTs with or without an enrichment design (in which patients are stabilized on the drug of interest during the open-label study, then randomized to receive the same drug or a comparator). Exclusion criteria were (1) studies with child/adolescent BD patients, (2) studies including patients with dual diagnosis of BD and other disorders, (3) continuation studies that randomly assigned patients with acute symptoms to treatment groups, and (4) studies of antidepressants.

Data synthesis and outcome measures

Primary outcome was recurrence rate of any mood episode at 6 months. Secondary outcomes were recurrence rates of depressive episodes and manic/hypomanic/mixed episodes and all-cause discontinuation rates at 6 months. Other outcomes were recurrence rates of any mood episodes, depressive episodes, and manic/hypomanic/mixed episodes and all-cause discontinuation rate at 1, 3, 9, 12, 18, and 24 months. The definitions of recurrence for each study are presented in online Supplementary Table S2. The results of data synthesis are summarized in online Supplementary Table S3.

Data extraction

The authors independently extracted data from all included studies. All analyses were based on intention-to-treat or modified intention-to-treat principles. When data required for the meta-analysis were incomplete, we contacted the original study investigators to obtain the unpublished data. However, we did not obtain any additional data by personal communication. We also searched for missing data in published systematic review articles. For studies with Kaplan–Meier survival curves, the recurrence rates and all-cause discontinuation rates were measured from the curves using a ruler.

Meta-analysis methods

We conducted a pair-wise meta-analysis to compare recurrence and discontinuation rates between the medication maintenance group and the medication discontinuation group (i.e. placebo group) using a random-effects model (DerSimonian & Laird, 1986). We calculated risk ratios (RRs) with 95% confidence intervals (95% CIs). When the ratio is statistically significant for a specific outcome, the number needed to treat to benefit/harm (NNTB/NNTH) was calculated as the reciprocal of the risk difference. We assessed the heterogeneity of the included studies using the I^2 statistics, with $I^2 \geqslant 50\%$ considered substantial heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003). We also conducted a single-group summary meta-analysis to calculate the exact event rates and 95% CIs for the primary and secondary outcomes in both the maintenance and discontinuation groups.

We conducted the first set of subgroup analyses to examine whether differences in the clinical characteristics of the assigned drugs influenced the effect size for the primary and secondary outcomes. The following comparisons were made: (1) LAI-SGAs *v.* oral medications, (2) LAI-SGAs *v.* oral SGAs (OSGAs), (3) LAI-SGAs *v.* lithium, (4) LAI-SGAs *v.* lamotrigine, (5) OSGAs *v.* lithium, (6) OSGAs *v.* lamotrigine, and (7) lithium *v.* lamotrigine. We also performed a second set of subgroup analyses on the exact event rates of the primary and secondary outcomes in the maintenance group using the same method as the first subgroup analysis. Although there were 15 enrichment studies, there were only two enrichment studies on lithium, which otherwise was the subject of the largest number of studies. All SGA studies [except for one olanzapine study (Vieta et al., 2012)] and all lamotrigine studies used an enriched design. Therefore, we did not perform subgroup analysis stratified by study design (enrichment *v.* without enrichment). Details of the first and second subgroup analyses are presented in online Supplementary appendix 1. In addition, we performed a third set of subgroup analyses to investigate whether differences in the characteristics of drugs used before randomization influenced the exact event rate of the primary and secondary outcomes in the discontinuation group. These analyses were also conducted using the same methods as in the first subgroup analyses. Furthermore, we performed a meta-regression analysis to evaluate the association of both primary and secondary outcomes with specific study features (the total number of patients, proportion of patients randomized to the placebo group, proportion of females, mean age, duration of the preliminary phase, and publication year).

We performed all statistical analyses using the Comprehensive Meta-Analysis Software Version 3 (Biostat Inc., Englewood, NJ, USA). Moreover, we corrected the results of the primary, subgroup, and meta-regression analyses for false discovery rate due to multiple comparisons using the Benjamini–Hochberg method (Benjamini & Hochberg, 1995). We assessed the methodological quality of the included studies according to the Cochrane Risk of Bias criteria (Cochrane Collaboration, http://www.cochrane.org/).

https://doi.org/10.1017/S0033291720003505 Published online by Cambridge University Press

Finally, we used funnel plots and Egger's regression tests to detect publication bias.

Results

Study characteristics

Online Supplementary Fig. S1 shows the literature search and selection strategy. The initial search retrieved 5469 articles, of which 2686 were eliminated as duplicates. Among the remaining articles, 2759 were eliminated based on review of the abstract and/or title. Again, we reviewed the full text of the remaining 24 articles, resulting in the elimination of 4 articles. We identified two additional DBRPCTs by a manual search through the reference lists of a review article (Miura et al., 2014). No further studies were found in the clinical trial registers. Finally, we identified 22 studies with a total of 5462 patients with BD (Amsterdam & Shults, 2010; Berwaerts, Melkote, Nuamah, & Lim, 2012; Bowden et al., 2000, 2003; Calabrese et al., 2000; Calabrese et al., 2003, 2017; Cundall, Brooks, & Murray, 1972; Dunner, Stallone, & Fieve, 1976; Fieve, Kumbaraci, & Dunner, 1976; Kane et al., 1982; Keck et al., 2007; Koyama et al., 2011; Melia, 1970; Prien, Caffey, & Klett, 1973; Prien, Klett, & Caffey, 1973; Quiroz et al., 2010; Szegedi et al., 2018; Tohen et al., 2006; Vieta et al., 2012; Weisler, Nolen, Neijber, Hellqvist, & Paulsson, 2011; Young et al., 2014) (53.78% female and mean age 40.37 ± 4.73 years). Online Supplementary Table S2 summarizes the characteristics of the included DBRPCTs. The mean study duration was 64.50 ± 69.35 weeks. The maintenance group included patients receiving aripiprazole (one study), asenapine (one study), divalproex (one study), LAI-aripiprazole (one study), LAI-risperidone (two studies), lamotrigine (four studies), lithium (12 studies), olanzapine (two studies), paliperidone (one study), and quetiapine (two studies). No study used a first-generation antipsychotic. Although there were 15 enrichment studies in total, there were only two enrichment studies on lithium. All SGA studies [except for one olanzapine study (Vieta et al., 2012)] and all lamotrigine studies used an enriched design. All studies other than one study (Bowden et al., 2000) employed the strategy of abrupt drug discontinuation for the placebo group. Twelve studies included only patients with bipolar I disorder, whereas one study included only patients with rapid cycling BD (Calabrese et al., 2000). Fifteen studies were industry sponsored. Most studies were of a high-quality design (online Supplementary Fig. S2).

Results of the meta-analyses

Primary and secondary outcomes

Figures 1–4 and online Supplementary Table S5 summarize the meta-analysis results for primary and secondary outcomes and the exact event rates of these outcomes. The recurrence rate of any mood episode at 6 months (the primary outcome) was significantly lower in the maintenance group than in the discontinuation group (RR = 0.61, 95% CI 0.54–0.70, $p = 0.000$, $I^2 = 75.33\%$, NNTB = 5, 95% CI 4–6). The maintenance group also exhibited significantly lower depressive episode recurrence rate (RR = 0.72, 95% CI 0.60–0.87, $p = 0.001$, $I^2 = 73.15\%$, NNTB = 13, 95% CI 8–29), manic/hypomanic/mixed episode recurrence rate (RR = 0.45, 95% CI 0.36–0.57, $p = 0.000$, $I^2 = 71.74\%$, NNTB = 6, 95% CI 5–9), and all-cause discontinuation rate (RR = 0.71, 95% CI 0.61–0.82, $p = 0.000$, $I^2 = 81.92\%$, NNTH = 6, 95% CI 4–8) than the discontinuation group at 6 months. We did not detect significant publication bias for primary and secondary outcomes (online Supplementary Fig. S3), but found considerable heterogeneity for the outcomes in the pair-wise meta-analyses.

Other outcomes

Figures 1–4 and online Supplementary Table S5 summarize the meta-analysis results for all other outcomes and exact event rates. The maintenance group demonstrated lower recurrence rates for any mood episode, depressive episodes, and manic/hypomanic/mixed episodes as well as lower all-cause discontinuation rate than the discontinuation group at 1, 3, 9, 12, 18, and 24 months.

Results of the subgroup analyses

Specific subgroups exhibited considerable heterogeneity.

Subgroup analyses for the primary outcome

Online Supplementary Table S6 summarizes the subgroup comparisons for the primary outcome. The recurrence rate of any mood episode at 6 months was significantly lower for all medications than that for placebo. The RRs were smaller for OSGA than for lithium and lamotrigine. The recurrence rate of depressive episodes at 6 months was also lower for oral medications, OSGAs, and lamotrigine than for placebo, and the RRs were lower for OSGAs than for LAI-SGAs and lithium. The recurrence rate of manic/hypomanic/mixed episodes at 6 months was lower for oral medication, LAI-SGAs, OSGAs, and lithium than for placebo, and the RRs were lower for OSGAs than for lithium and lamotrigine. Compared with placebo, oral medications, LAI-SGAs, OSGAs, and lamotrigine were associated with lower all-cause discontinuation at 6 months. The RR for OSGAs was lower than those for LAI-SGAs, lithium, and lamotrigine.

Single-group summary meta-analysis of the maintenance group

Online Supplementary Fig. S4 and Table S6 summarize the maintenance subgroup analyses.

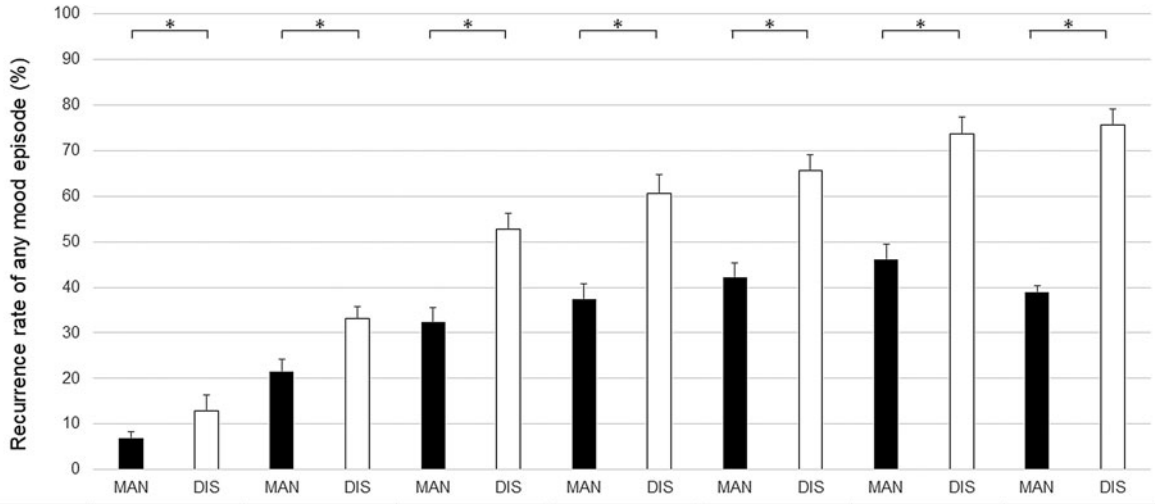
The recurrence rates of any mood episode and depressive episodes as well as all-cause discontinuation at 6 months were lower for LAI-SGAs and OSGAs than for lithium and lamotrigine. The recurrence rate of manic/hypomanic/mixed episodes was also lower for OSGAs than for lithium.

Single-group summary meta-analysis of the discontinuation group

Figure 5 and online Supplementary Table S7 summarize the results of the discontinuation subgroup analyses.

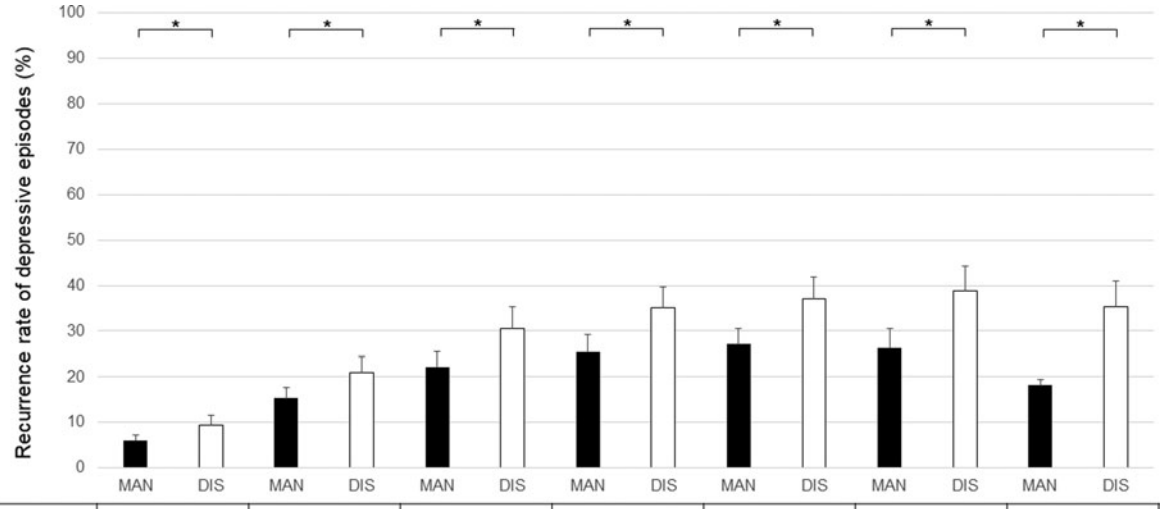
The recurrence rate of any mood episode at 6 months was lower for the SGA-LAI discontinuation subgroup than for the oral medication and lamotrigine discontinuation subgroups. Similarly, the rate of depressive episode recurrence at 6 months was lower in the SGA-LAI discontinuation subgroup than in the oral medication, OSGA, and lamotrigine discontinuation subgroups. In contrast, we did not detect any subgroup differences in manic/hypomanic/mixed episode recurrence rate at 6 months among the discontinuation subgroups. However, all-cause discontinuation at 6 months was lower in the SGA-LAI and OSGA discontinuation subgroups than in the lamotrigine discontinuation subgroup.

https://doi.org/10.1017/S0033291720003505 Published online by Cambridge University Press



	1 m	3 m	6 m	9 m	12 m	18 m	24 m
N (n)	19 (4165)	20 (4196)	20 (4178)	16 (3628)	16 (3628)	12 (2271)	8 (1771)
RR (95% CI)	0.59 (0.42, 0.83)	0.63 (0.51, 0.77)	0.61 (0.54, 0.70)	0.60 (0.51, 0.70)	0.63 (0.55, 0.72)	0.61 (0.53, 0.72)	0.49 (0.45, 0.52)
NNTB (95% CI)	20 (12, 50)	9 (6, 15)	5 (4, 6)	4 (3, 6)	4 (3, 6)	3 (3, 5)	2 (2, 3)

Fig. 1. Recurrence rate of any mood episode. *Adjusted *p* after false discovery rate correction (Benjamini–Hochberg method) <0.05. Error bar represents standard error. 95% CI, 95% confidence interval; DIS, discontinuation group; m, month(s); MAN, maintenance group; N, number of comparisons; *n*, number of patients; NNTB, number needed to treat to benefit; RR, risk ratio.



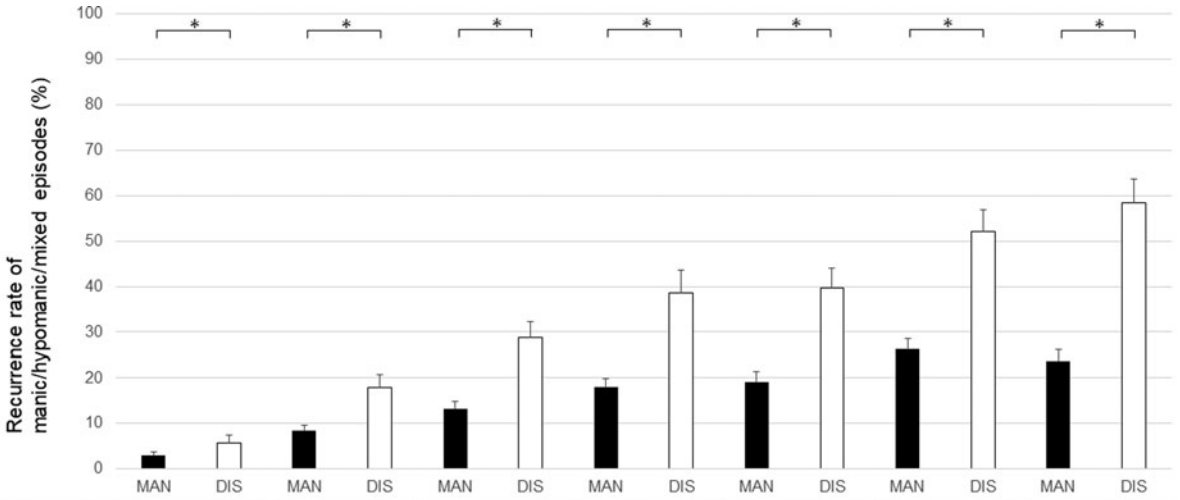
	1 m	3 m	6 m	9 m	12 m	18 m	24 m
N (n)	18 (3770)	18 (3770)	18 (3770)	16 (3415)	17 (3680)	13 (2324)	6 (1555)
RR (95% CI)	0.61 (0.44, 0.85)	0.71 (0.59, 0.86)	0.72 (0.60, 0.87)	0.71 (0.58, 0.87)	0.71 (0.59, 0.84)	0.65 (0.50, 0.84)	0.47 (0.39, 0.58)
NNTB (95% CI)	29 (18, 77)	17 (11, 34)	13 (8, 29)	11 (7, 25)	9 (6, 18)	7 (5, 16)	5 (4, 9)

Fig. 2. Recurrence rate of depressive episodes. *Adjusted *p* after false discovery rate correction (Benjamini–Hochberg method) <0.05. Error bar represents standard error. 95% CI, 95% confidence interval; DIS, discontinuation group; m, month(s); MAN, maintenance group; N, number of comparisons; n, number of patients; NNTB, number needed to treat to benefit; RR, risk ratio.

Results of meta-regression analysis

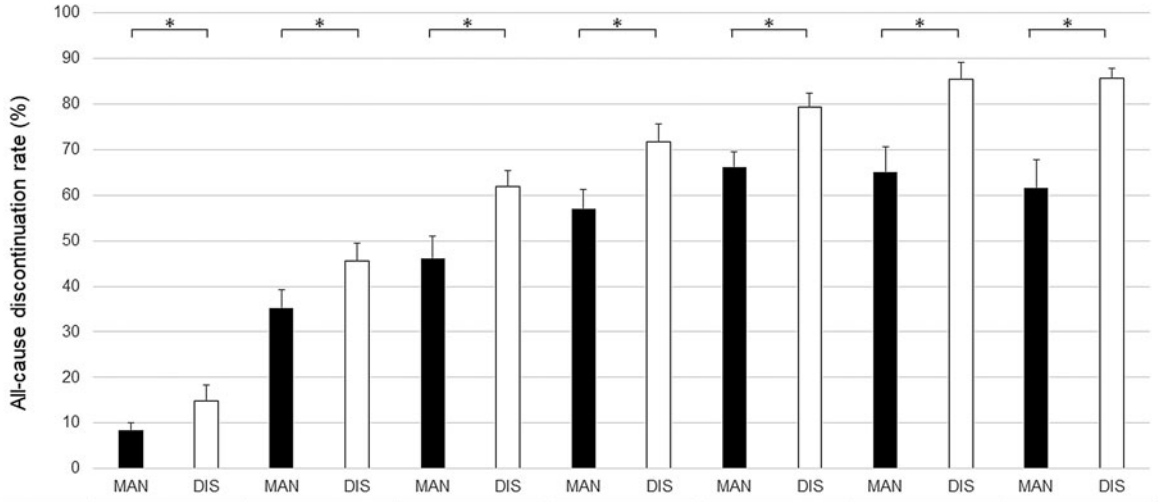
Online Supplementary Table S8 summarizes the results of meta-regression analysis. We found no associations between the specified study features and effect sizes for the primary and secondary outcomes. However, the publication year was associated with the event rates of all efficacy outcomes in the maintenance

https://doi.org/10.1017/S0033291720003505 Published online by Cambridge University Press



	1 m	3 m	6 m	9 m	12 m	18 m	24 m
N (n)	17 (3717)	17 (3717)	17 (3717)	15 (3362)	17 (3680)	13 (2324)	6 (1555)
RR (95% CI)	0.49 (0.29, 0.83)	0.44 (0.34, 0.56)	0.45 (0.36, 0.57)	0.46 (0.37, 0.58)	0.47 (0.39, 0.58)	0.49 (0.41, 0.60)	0.39 (0.35, 0.44)
NNTB (95% CI)	34 (21, 91)	9 (7, 13)	6 (5, 9)	5 (4, 7)	5 (4, 7)	4 (3, 5)	3 (2, 3)

Fig. 3. Recurrence rate of manic/hypomanic/mixed episodes. *Adjusted *p* after false discovery rate correction (Benjamini–Hochberg method) <0.05. Error bar represents standard error. 95% CI, 95% confidence interval; DIS, discontinuation group; m, month(s); MAN, maintenance group; N, number of comparisons; n, number of patients; NNTB, number needed to treat to benefit; RR, risk ratio.



	1 m	3 m	6 m	9 m	12 m	18 m	24 m
N (n)	11 (2640)	11 (2640)	13 (2906)	9 (2355)	13 (2954)	9 (1770)	7 (1771)
RR (95% CI)	0.60 (0.45, 0.80)	0.75 (0.65, 0.85)	0.71 (0.61, 0.82)	0.76 (0.67, 0.86)	0.82 (0.75, 0.88)	0.72 (0.61, 0.86)	0.68 (0.53, 0.86)
NNTH (95% CI)	19 (12, 42)	9 (6, 13)	6 (4, 8)	6 (4, 9)	7 (5, 11)	5 (3, 9)	4 (3, 10)

Fig. 4. All-cause discontinuation rate. *Adjusted *p* after false discovery rate correction (Benjamini–Hochberg method) <0.05. Error bar represents standard error. 95% CI, 95% confidence interval; DIS, discontinuation group; m, month(s); MAN, maintenance group; N, number of comparisons; n, number of patients; NNTH, number needed to treat to harm; RR, risk ratio.

https://doi.org/10.1017/S0033291720003505 Published online by Cambridge University Press

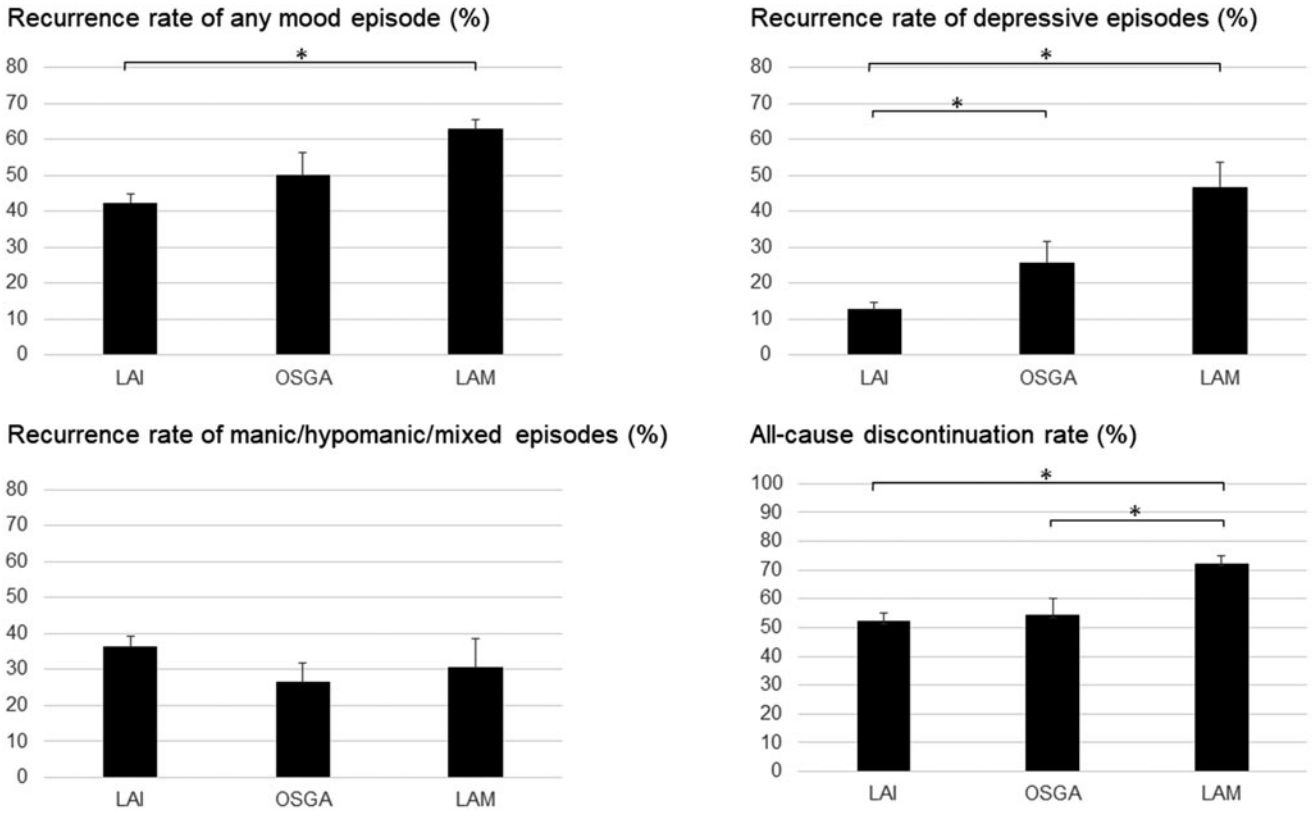


Fig. 5. Pooled event rates in the discontinuation subgroups stratified by drugs used before randomization. *Adjusted *p* after false discovery rate correction (Benjamini–Hochberg method) <0.05. Error bar represents standard error. LAI, long-acting injection-second generation antipsychotics; LAM, lamotrigine; OSGA, oral second generation antipsychotics.

group. Moreover, the publication year was associated with the recurrence rate of any mood episode and all-cause discontinuation rate in the discontinuation group.

Discussion

To the best of our knowledge, this is the first systematic review and meta-analysis of DBRPCTs comparing recurrence rates between stable BD patients maintaining SGA or MS treatment and those discontinuing medication. The recurrence risks for any mood episode and manic/hypomanic/mixed episodes and the risk of all-cause discontinuation increased by 1 month after discontinuation, and all these risks were very high after at least 3 months of antipsychotic or MS discontinuation with large effect sizes. In contrast, the depressive episode recurrence rate increased more gradually with the duration of discontinuation. Thus, we conclude that the maintenance of antipsychotic or MS therapy is beneficial for preventing recurrence over the first 24 months among clinically stable BD patients. Notably, the risk of manic/hypomanic/mixed episode recurrence increased immediately after discontinuation, whereas the risk of depressive recurrence was delayed. However, some patients with clinically stable BD did not suffer a recurrence after discontinuation; therefore, it is critical to identify clinical measures or biomarkers predictive of this subpopulation for improved clinical decision-making and tailoring of personalized treatment.

Subgroup analysis revealed no significant difference in any outcome between patients receiving different drug formulations (i.e. LAI-SGAs *v.* oral medication). However, there was considerable heterogeneity for all outcomes in the oral medication subgroup. We suspect that this heterogeneity stems from pooling of oral drugs (OSGAs, lithium, and lamotrigine) with different efficacies and tolerability. Therefore, we can draw no specific conclusions regarding the safe (recurrence-free) discontinuation of pooled oral medications.

Subgroup analyses revealed that LAI-SGA and OSGA were superior to placebo for preventing the recurrence of any mood episode and manic/hypomanic/mixed episodes, with no significant differences in effect sizes between LAI-SGA and OSGA subgroups. In contrast, there is a significant difference in effect sizes for recurrence rate of depressive episodes between LAI-SGA and OSGA subgroups, even though the average rate of depressive episode recurrence was similar between LAI-SGA and OSGA subgroups. However, the exact event rate of depressive episode recurrence was lower in the LAI-SGA discontinuation subgroup than in the OSGA discontinuation subgroup. Therefore, the lack of a significant difference in depressive episode recurrence rate between LAI-SGA and placebo may be owing to the low recurrence rate of depressive episodes in the LAI-SGA discontinuation subgroup (i.e. placebo group). Although current guidelines do not recommend LAI-SGAs for the prevention of depressive episodes (Yatham et al., 2018), based on our findings, we suggest that LAI-SGAs may also prevent depressive episode recurrence in BD patients. Systemic drug concentration decreases gradually after LAI-SGA discontinuation according to half-life measurements (online Supplementary Table S1). Therefore, the small

https://doi.org/10.1017/S0033291720003505 Published online by Cambridge University Press

residual amount remaining may be sufficient to prevent depressive episode recurrence. However, the RR for all-cause discontinuation rate was higher in patients receiving LAI-SGAs than in those receiving OSGAs. The exact event rate was higher in patients receiving LAI-SGAs before adjustment for multiple testing. Thus, LAI-SGAs might be inferior to OSGAs in acceptability. However, very few studies included LAI-SGA subgroup. These subgroup analyses did not compare LAI-SGAs with identical oral formulations of SGAs. Therefore, we recommend larger studies comparing LAI-SGAs with identical oral formulations of SGAs in order to confirm our results.

In addition, subgroup analysis showed that lithium prevents recurrence of any mood episode and manic/hypomanic/mixed episodes but not depressive episodes. There was no difference in all-cause discontinuation between lithium and placebo, but exact all-cause discontinuation was higher than that for LAI-SGA or OSGA treatment. Although lithium appears not to have a good risk benefit–balance for BD compared with LAI-SGAs and OSGAs, its efficacy and tolerability may be underestimated because only 2 of 12 lithium studies used an enrichment design. The Finnish nationwide cohort of 18 018 patients with BD (mean follow-up time = 7.2 years) demonstrated that lithium and LAI-antipsychotics were the most effective at preventing hospitalization due to mental or physical illness compared with no drug use (Lahteenvuo et al., 2018). Thus, there appear to be inconsistencies between the results of our meta-analysis, which included randomized controlled trials (RCTs) (providing the most robust evidence), and those of the cohort study (reflecting 'real-world' routine clinical practice). We could not simply compare results between the studies for the following reasons (Blonde, Khunti, Harris, Meizinger, & Skolnik, 2018; Lahteenvuo et al., 2018). First, the study durations of RCTs are generally shorter than those of non-RCT studies. Second, the symptoms of trial populations are evaluated in more detail than those of patient populations in clinical practice. Hence, symptoms may be detected earlier and early intervention given to trial populations than in clinical practice. Third, because RCTs often have stringent inclusion and exclusion criteria (e.g. excluding patients with the most comorbidities and the highest severity of illness such as suicidal ideation and suicidal attempt), trial populations are often not representative of those in clinical practice. Thus, although RCTs are a well-established methodology for gathering robust evidence for safety and efficacy of medical interventions, it might be difficult for RCTs to provide results that reflect the effectiveness of treatments in actual clinical settings.

In contrast to lithium, lamotrigine prevented recurrence of any mood episode and depressive episodes but not manic/hypomanic/mixed episodes. However, patients receiving lamotrigine exhibited a higher exact event rate of depressive episodes than those receiving LAI-SGAs and OSGAs. Moreover, because the exact recurrence rate of depressive episodes in the lamotrigine discontinuation subgroup was higher than that in the LAI-SGA discontinuation subgroup (there was also a difference in event rate between the OSGA discontinued subgroup and lamotrigine discontinued subgroup before adjusting for multiple testing), we suggest that abrupt discontinuation of lamotrigine may increase the recurrence rate of depressive episodes. Therefore, recent guideline recommend lamotrigine for the prevention of depressive episodes (Yatham et al., 2018), but we consider that SGAs is preferred over lamotrigine. However, one lamotrigine study included only rapid cycling BD patients (Calabrese et al., 2000). We conducted an additional subgroup analysis excluding this study but the results were similar to those of the original subgroup analysis (online Supplementary Table S6 and S7).

Our study had several limitations. First, we found considerable heterogeneity for primary and secondary outcomes in the pair-wise meta-analysis, possibly because the included studies used different definitions of recurrence (online Supplementary Table S2). There are two methods to discontinue drugs: abrupt and gradual. This difference might have influenced the results. There was only one study using abrupt discontinuation strategy (Bowden et al., 2000). Although we performed additional subgroup analysis excluding this study for primary and secondary outcomes, the results of this subgroup analysis were similar to the results of the original pair-wise meta-analysis (online Supplementary Table S9). Other potential reason could be that the OSGA subgroup included multiple OSGAs with distinct efficacies and tolerability. Moreover, the heterogeneity in the LAI-SGA subgroup may have stemmed from a small sample size as only three studies included this drug type. The heterogeneity of the lithium subgroup could result from the inclusion of both enrichment and non-enrichment design studies. However, there was only one enrichment study for the subgroup analysis (Cundall et al., 1972). Due to this, we could not examine the difference. Another potential source of heterogeneity is the publication year as our meta-regression analysis revealed that newer studies reported lower event rates for some outcomes in both maintenance and discontinuation groups. These did not influence the effect sizes for the primary and secondary outcomes of the primary meta-analyses. We suggest that factors affecting maintenance-response and discontinuation-response act in the same direction and to a similar extent, and thus have no net influence on the effect size. Second, the observation period in our study was only 2 years. Thus, the long-term efficacy and safety of drugs still need to be verified. Third, we did not cover important clinical issues that might inform treatment decision-making in routine clinical practice (e.g. combination with non-pharmacological treatments).

In conclusion, our results suggest that maintaining SGAs or MSs is beneficial for preventing recurrences over at least 2 years among clinically stable BD patients. Notably, the discontinuation of medications for ⩾1 month significantly increased the risk of recurrence. However, some patients with clinically stable BD did not suffer recurrence after discontinuation. It is therefore critical to identify markers predictive of recurrence risk following discontinuation of medication for BD.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291720003505

Data. Data used for the current study were reported in the articles of the studies included in our meta-analysis.

Acknowledgements. We would like to thank all participants of the current study.

Author contributions. Dr Kishi had full access to all data and takes full responsibility for the integrity of the data along with the accuracy of the data analysis. Dr Kishi was involved in the study concept and design, and performed the statistical analyses. Drs. Kishi, Matsuda, Sakuma, and Okuya acquired and interpreted the data. All authors wrote the manuscript. Drs. Iwata and Mishima supervised the review.

Financial support. The present study was supported by a Health and Labor Sciences Research Grant (H29-Seishin-Ippan-001, 19GC1012).

Conflict of interest. The authors declare no conflicts of interest relating to the subject of this study. Interests from the past three years are as follows. Dr Kishi received speaker's honoraria from Daiichi Sankyo, Dainippon Sumitomo, Eisai, Janssen, Otsuka, Meiji, Mochida, MSD, and Tanabe-Mitsubishi (Yoshitomi), as well as a research grant from the Japanese Ministry of Health, Labour and Welfare (H29-Seishin-Ippan-001, 19GC1012), a Grant-in-Aid for Scientific Research (C, 19K08082), and a grant from the Fujita Health University School of Medicine. Dr Matsuda has received speaker's honoraria from Dainippon Sumitomo, Janssen, Kyowa, Otsuka, Tanabe-Mitsubishi, and Yoshitomi. Dr Sakuma has received speaker's honoraria from Eisai, Kissei, Meiji, Otsuka, and Torii, a Fujita Health University School of Medicine research grant, and a Grant-in-Aid for Young Scientists (B). Dr Okuya has received a speaker's honoraria from Meiji. Dr Mishima has received research support from the Japanese Ministry of Health, Labour and Welfare (H29-Seishin-Ippan-001, 19GC1012) and the Japanese Ministry of Education, Culture, Sports, Science and Technology Collaborative research fund with Taisho, speaker's honoraria from Eisai, MSD, Takeda, Astellas, Pfizer, Otsuka, Mochida, Mitsubishi Tanabe, Yoshitomi, and Janssen, and research grants from Eisai, Nobelpharma, Otsuka, and Takeda. Dr Iwata has received speaker's honoraria from Astellas, Dainippon Sumitomo, Eli Lilly, GlaxoSmithKline, Janssen, Yoshitomi, Otsuka, Meiji, Shionogi, Novartis, and Pfizer as well as research grants from Eisai, Takeda, Dainippon Sumitomo, and Otsuka.

References

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. *The American Journal of Psychiatry*, *167*(7), 792–800. doi: 10.1176/appi.ajp.2009.09020284.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Berwaerts, J., Melkote, R., Nuamah, I., & Lim, P. (2012). A randomized, placebo- and active-controlled study of paliperidone extended-release as maintenance treatment in patients with bipolar I disorder after an acute manic or mixed episode. *Journal of Affective Disorders*, *138*(3), 247–258. doi: 10.1016/j.jad.2012.01.047.

Blonde, L., Khunti, K., Harris, S. B., Meizinger, C., & Skolnik, N. S. (2018). Interpretation and impact of real-world clinical data for the practicing clinician. *Advances in Therapy*, *35*(11), 1763–1774. doi: 10.1007/s12325-018-0805-y.

Bowden, C. L., Calabrese, J. R., McElroy, S. L., Gyulai, L., Wassef, A., Petty, F., … Wozniak, P. J. (2000). A randomized, placebo-controlled 12-month trial of divalproex and lithium in treatment of outpatients with bipolar I disorder. Divalproex Maintenance Study Group. *Archives of General Psychiatry*, *57*(5), 481–489. doi: 10.1001/archpsyc.57.5.481.

Bowden, C. L., Calabrese, J. R., Sachs, G., Yatham, L. N., Asghar, S. A., Hompland, M., … Lamictal 606 Study, G. (2003). A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I disorder. *Archives of General Psychiatry*, *60*(4), 392–400. doi:10.1001/archpsyc.60.4.392.

Calabrese, J. R., Bowden, C. L., Sachs, G., Yatham, L. N., Behnke, K., Mehtonen, O. P., … Lamictal 605 Study, G. (2003). A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently depressed patients with bipolar I disorder. *The Journal of Clinical Psychiatry*, *64*(9), 1013–1024. doi:10.4088/jcp.v64n0906.

Calabrese, J. R., Sanchez, R., Jin, N., Amatniek, J., Cox, K., Johnson, B., … Carson, W. H. (2017). Efficacy and safety of aripiprazole once-monthly in the maintenance treatment of bipolar I disorder: A double-blind, placebo-controlled, 52-week randomized withdrawal study. *The Journal of Clinical Psychiatry*, *78*(3), 324–331. doi: 10.4088/JCP.16m11201.

Calabrese, J. R., Suppes, T., Bowden, C. L., Sachs, G. S., Swann, A. C., McElroy, S. L., … Monaghan, E. T. (2000). A double-blind, placebo-controlled, prophylaxis study of lamotrigine in rapid-cycling bipolar disorder. Lamictal 614 Study Group. *The Journal of Clinical Psychiatry*, *61*(11), 841–850. doi: 10.4088/jcp.v61n1106.

Cundall, R. L., Brooks, P. W., & Murray, L. G. (1972). A controlled evaluation of lithium prophylaxis in affective disorders. *Psychological Medicine*, *2*(3), 308–311. doi: 10.1017/s0033291700042616.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177–188.

Dunner, D. L., Stallone, F., & Fieve, R. R. (1976). Lithium carbonate and affective disorders. V: A double-blind study of prophylaxis of depression in bipolar illness. *Archives of General Psychiatry*, *33*(1), 117–120. doi: 10.1001/archpsyc.1976.01770010073014.

Fieve, R. R., Kumbaraci, T., & Dunner, D. L. (1976). Lithium prophylaxis of depression in bipolar I, bipolar II, and unipolar patients. *The American Journal of Psychiatry*, *133*(8), 925–929. doi: 10.1176/ajp.133.8.925.

Grande, I., Berk, M., Birmaher, B., & Vieta, E. (2016). Bipolar disorder. *Lancet (London, England)*, *387*(10027), 1561–1572. doi: 10.1016/S0140-6736(15)00241-X.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, *327*(7414), 557–560. doi: 10.1136/bmj.327.7414.557.

Kane, J. M., Quitkin, F. M., Rifkin, A., Ramos-Lorenzi, J. R., Nayak, D. D., & Howard, A. (1982). Lithium carbonate and imipramine in the prophylaxis of unipolar and bipolar II illness: A prospective, placebo-controlled comparison. *Archives of General Psychiatry*, *39*(9), 1065–1069. doi: 10.1001/archpsyc.1982.04290090053011.

Keck, P. E., Jr., Calabrese, J. R., McIntyre, R. S., McQuade, R. D., Carson, W. H., Eudicone, J. M., … Aripiprazole Study, G. (2007). Aripiprazole monotherapy for maintenance therapy in bipolar I disorder: A 100-week, double-blind study versus placebo. *The Journal of clinical psychiatry*, *68*(10), 1480–1491. doi:10.4088/jcp.v68n1003.

Kishi, T., Oya, K., & Iwata, N. (2016). Long-acting injectable antipsychotics for prevention of relapse in bipolar disorder: A systematic review and meta-analyses of randomized controlled trials. *The International Journal of Neuropsychopharmacology*, *19*(9), pyw038. doi: 10.1093/ijnp/pyw038.

Koyama, T., Higuchi, T., Yamawaki, S., Kanba, S., Terao, T., & Shinohara, A. (2011). Study SCA104779, an evaluation of BW430C (lamotrigine) versus placebo in the prevention of mood episodes in bipolar I disorder patients. *Japanese Journal of Clinical Psychiatry*, *40*(3), 369–383.

Lahteenvuo, M., Tanskanen, A., Taipale, H., Hoti, F., Vattulainen, P., Vieta, E., & Tiihonen, J. (2018). Real-world effectiveness of pharmacologic treatments for the prevention of rehospitalization in a finnish nationwide cohort of patients with bipolar disorder. *JAMA Psychiatry*, *75*(4), 347–355. doi: 10.1001/jamapsychiatry.2017.4711.

Melia, P. I. (1970). Prophylactic lithium: A double-blind trial in recurrent affective disorders. *The British Journal of Psychiatry: The Journal of Mental Science*, *116*(535), 621–624. doi: 10.1192/bjp.116.535.621.

Miura, T., Noma, H., Furukawa, T. A., Mitsuyasu, H., Tanaka, S., Stockton, S., … Kanba, S. (2014). Comparative efficacy and tolerability of pharmacological treatments in the maintenance treatment of bipolar disorder: A systematic review and network meta-analysis. *The Lancet. Psychiatry*, *1*(5), 351–359. doi: 10.1016/S2215-0366(14)70314-1.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ: British Medical Journal*, *339*, b2535. doi: 10.1136/bmj.b2535.

Prien, R. F., Caffey, E. M., Jr., & Klett, C. J. (1973). Prophylactic efficacy of lithium carbonate in manic-depressive illness. Report of the Veterans Administration and National Institute of Mental Health collaborative study group. *Archives of General Psychiatry*, *28*(3), 337–341. doi:10.1001/archpsyc.1973.01750330035006.

Prien, R. F., Klett, C. J., & Caffey, E. M., Jr. (1973). Lithium carbonate and imipramine in prevention of affective episodes. A comparison in recurrent affective illness. *Archives of General Psychiatry*, *29*(3), 420–425. doi:10.1001/archpsyc.1973.04200030104017.

Quiroz, J. A., Yatham, L. N., Palumbo, J. M., Karcher, K., Kushner, S., & Kusumakar, V. (2010). Risperidone long-acting injectable monotherapy in the maintenance treatment of bipolar I disorder. *Biological Psychiatry*, *68*(2), 156–162. doi: 10.1016/j.biopsych.2010.01.015.

Szegedi, A., Durgam, S., Mackle, M., Yu, S. Y., Wu, X., Mathews, M., & Landbloom, R. P. (2018). Randomized, double-blind, placebo-controlled

trial of asenapine maintenance therapy in adults with an acute manic or mixed episode associated with bipolar I disorder. *The American Journal of Psychiatry*, *175*(1), 71–79. doi: 10.1176/appi.ajp.2017.16040419.

Tohen, M., Calabrese, J. R., Sachs, G. S., Banov, M. D., Detke, H. C., Risser, R., … Bowden, C. L. (2006). Randomized, placebo-controlled trial of olanzapine as maintenance therapy in patients with bipolar I disorder responding to acute treatment with olanzapine. *The American Journal of Psychiatry*, *163* (2), 247–256. doi: 10.1176/appi.ajp.163.2.247.

Vieta, E., Montgomery, S., Sulaiman, A. H., Cordoba, R., Huberlant, B., Martinez, L., & Schreiner, A. (2012). A randomized, double-blind, placebo-controlled trial to assess prevention of mood episodes with risperidone long-acting injectable in patients with bipolar I disorder. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, *22*(11), 825–835. doi: 10.1016/j.euroneuro.2012.03.004.

Weisler, R. H., Nolen, W. A., Neijber, A., Hellqvist, A., Paulsson, B., & Trial 144 Study, I. (2011). Continuation of quetiapine versus switching to placebo or lithium for maintenance treatment of bipolar I disorder (trial 144: A randomized controlled study). *The Journal of Clinical Psychiatry*, *72*(11), 1452–1464. doi:10.4088/JCP.11m06878.

Yatham, L. N., Kennedy, S. H., Parikh, S. V., Schaffer, A., Bond, D. J., Frey, B. N., … Berk, M. (2018). Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. *Bipolar Disorders*, *20*(2), 97–170. doi: 10.1111/bdi.12609.

Young, A. H., McElroy, S. L., Olausson, B., Paulsson, B., Embolden, I., & Embolden, I. I. I. (2014). A randomised, placebo-controlled 52-week trial of continued quetiapine treatment in recently depressed patients with bipolar I and bipolar II disorder. *The World Journal of Biological Psychiatry*, *15*(2), 96–112. doi: 10.3109/15622975.2012.665177.