COMPARATIVE EFFICACY AND ACCEPTABILITY OF PHARMACOLOGICAL TREATMENTS IN THE ACUTE PHASE TREATMENT OF BIPOLAR DEPRESSION: A MULTIPLE TREATMENT META-ANALYSIS [PROTOCOL]

Tomofumi Miura, Toshi A. Furukawa, Hisashi Noma, Shiro Tanaka, Keisuke Motomura, Hiroshi Mitsuyasu, Satomi Katsuki, Andrea Cipriani, Sarah Stockton, John Geddes, Georgia Salanti, Shigenobu Kanba

Background

Description of the condition

In late 19th century, Emil Kraepelin classified the major endogenous psychoses into two psychotic illnesses, manic-depressive illness and dementia praecox (Kraepelin's dichotomy) (Trede et al. , 2005). His concept of manic-depressive illness had covered all of the major clinical forms of severe, moderate and mild melancholia. In 1957, Leonhard proposed to distinguish Kraepelin's manic-depressive illness into bipolar illness, in which patients had histories of both depression and mania, and monopolar illness, in which patients only had histories of depression (Leonhard, 1979). This basic idea of the bipolar-unipolar distinction was supported by Angst and by Perris in 1966 (Angst, 1966, Perris, 1966), and it continues in the recent diagnostic classification system of mood disorders in DSM-IV(American Psychiatric Association, 1994) and ICD-10 (WHO, 1993). Bipolar disorder is a complex disorder, which is characterized by recurrent episodes of depression and mania (bipolar I disorder) or hypomania (bipolar II disorder)(American Psychiatric Association, 1994). The lifetime prevalence of any bipolar disorders, bipolar I and II disorders have been estimated at 1.1%, 0.7% and 0.5% respectively using the World Mental Health Survey version of the WHO Composite International Diagnostic Interview (Suppes et al. , 2001).

The mean age at onset of bipolar disorder is reported to be in the early 20s, but its complex clinical features make its diagnosis difficult and there is a difference of about 8 years between age at onset and age at first treatment (Suppes, Leverich, 2001). Moreover, bipolar disorder has a chronic course of illness. The long-term prospective follow-up studies revealed that the percentages of bipolar I patient who remained in remission for years are substantially low, 28% for 4 years and about 10% for 5 years (Goodwin and Jamison, 2007, Keller et al. , 1993, Tohen et al. , 1990).

Bipolar disorder is associated with lower health-related quality of life, lower social functioning, unemployment and lower productivity than the general non-ill population (Dean et al. , 2004). Altogether, bipolar disorder is estimated to be the 30th leading cause of disability-adjusted life years lost for the human kind according to the latest WHO Global Burden of Disease study (WHO, 2008).

The long-term follow-up natural history studies for bipolar disorder show that the amounts of time periods that bipolar I and bipolar II patient have been in depressive episode were estimated at 31.9% and 50.3%, respectively (Judd et al. , 2003, Judd et al. , 2002). The impact of the depressive episodes on the course of bipolar disorder and on the social disability of the patient is revealed to be greater than those of manic episodes (Calabrese et al. , 2004).

Description of the intervention

The differences in pharmacological treatment between bipolar depression and unipolar depression have not received much attention for many years. Clinical guidelines recommended using anti-manic agents alone or in combination with some other class of drugs for treatment in acute bipolar depression [National Institute for Health and Clinical Excellence (NICE) Bipolar disorder: the management of bipolar disorder in adults, children and adolescents, in primary and secondary care (Issue date: July 2006). Available at www.nice.org.uk/CG038.]. Lithium, one of the oldest psychotropic agents that were proved to have anti-manic activity in 1949 (Cade, 1949), has been used not only for maintenance treatment in bipolar disorder but for a treatment in acute depressive episode. Not all but some anti-epileptics (valproic acid, carbamazepine) are anti-manic drugs (Cipriani et al. , 2011) and are used for the treatment in acute bipolar depression. Lamotrigine is different from other anti-epileptics, which does not have anti-manic activity (Cipriani, Barbui, 2011) but may have some efficacy on acute bipolar depression (Geddes et al. , 2009).

Antidepressants have been used for the treatment of acute bipolar depression in most cases in combination with anti-manic agents. However, they are supposed to be associated with risks of switching to manic/hypomanic episode and with those of cycle acceleration in bipolar patients, especially in bipolar I disorder. The Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD), a large pragmatic clinical trial sponsored by the National Institute of Mental Health (NIMH), has demonstrated that a mood stabilizer plus adjunctive antidepressant (paroxetine or bupropion) therapy was not superior in efficacy than one plus placebo and that rates of treatment-associated mood switching were not different among two arms (Sachs et al. , 2007). The efficacy and acceptability of antidepressants in acute bipolar depression do not yet reach the consensus conclusion among psychiatrist.

While anti-psychotics have been typically used for the treatment of acute mania, recent studies have demonstrated that some second-generation antipsychotics, alone or in combination with antidepressant, also

were efficacious in acute bipolar depression (Vieta et al. , 2010).

Why it is important to do this review

Among the various treatment options available, it is clinically very important to know what pharmacological intervention, alone or in combination, is more efficacious and acceptable than others for the treatment in acute bipolar depression. Many clinical studies have been conducted to investigate the treatment effect of new pharmacological strategies and several systematic reviews and meta-analyses have already been published in the literature (Chiesa et al. , 2012, Geddes, Calabrese, 2009, Sidor and Macqueen, 2011, Vieta, Locklear, 2010).
Multiple treatment meta-analysis (MTM, also known as network meta-analysis) is a recently developed statistical technique that can integrate both direct and indirect comparison of the effects (Higgins and Whitehead, 1996, Lumley, 2002). MTM is considered to have two major advantages compared with conventional pair-wise methods, one is to increase precision of the relative efficacy between two treatments with combining both direct and indirect evidences, and another is to facilitate simultaneous inference regarding all treatments, that means relative efficacy of the two treatments can be estimated even when there is no study that directly compare the efficacy between two interventions.
MTM has already been used successfully in psychiatry (Cipriani, Barbui, 2011, Cipriani et al. , 2009) and in other fields of medicine (Psaty et al. , 2003, Stettler et al. , 2007). Considering how important comparative efficacy could be for clinical practice and policy making, it is useful to use all the available evidence to estimate potential differences in efficacy among treatments. We are therefore in dire need of a comparative effectiveness research with MTM including such newer pharmacological treatment of acute bipolar depression to inform our clinical practices.

Objectives

To compare the efficacy and acceptability of different drugs or combinations of drugs in the acute phase treatment of bipolar depression in adults.

Methods

Criteria for considering studies for this review

Types of studies

Inclusion criteria
Double-blind randomized controlled trials (RCTs) comparing any psychotropic agent with placebo, or one another in the treatment of acute major depressive episode in bipolar I disorder or bipolar II disorder will be included. Studies which include both unipolar and bipolar participants will be accepted if data are available for bipolar participants separately. We will accept studies that focused on specific conditions like rapid cycling.
Exclusion criteria
Quasi-randomized controlled trials, in which treatment assignment is decided through methods such as alternate days of the week, will be excluded. We will exclude open-label RCTs.

Types of participants

Participant characteristics

Inclusion criteria
Participants aged 18 or older, of both sexes with a primary diagnosis of acute major depressive episode in bipolar I disorder or bipolar II disorder, diagnosed according to any of the following operationalized criteria: Research Diagnostic Criteria, DSM-III, DSM-III-R, DSM-IV, DSM-IV-TR or ICD-10. Operationalized criteria essentially resembling these official ones will also be eligible.
Exclusion criteria
Bipolar disorder in children and adolescents is difficult to diagnosis because of its atypical symptoms. It also occurs with common child-onset mental disorders, including attention deficit/hyperactivity disorder (ADHD). We will exclude childhood bipolar disorder because special considerations are needed for its pharmacological treatment.

Setting

Inclusion criteria

Studies conducted in both in inpatient and outpatient settings will be included.

Types of interventions

Included interventions

We will include all the pharmacological interventions with prescription drugs in the treatment of acute major depressive episode in bipolar I, or II disorder, even if they are not licensed in any countries.

Excluded interventions

We will exclude the interventions with over-the-counter drugs, herbal medicine or nutritional supplement. Psychological therapy will not be the focus for this review. However, if the same type and amount of psychosocial intervention is provided to two arms which compared two drug treatments, such studies will be included.

Outcome measures

Primary outcomes

(1) Treatment efficacy: number of patients who respond to treatment, based on changes on Montgomery-Asberg Depression Rating Scale (MADRS) (Montgomery 1979) or Hamilton Rating Scale for Depression (HAM-D) (Hamilton 1960), or any other validated depression scale at the end of acute phase treatment (8 weeks, range 4-12 weeks). Many studies define response by 50% or greater reduction on the rating scale; we will accept the study authors' original definition. If the original authors report several outcomes corresponding with our definition of response, we will give preference to MADRS. Any version of HAM-D will be accepted.
(2) Treatment acceptability: number of participants who drop out of treatment for any reasons during the first 8 weeks of treatment (range: 4-12 weeks).

Secondary outcomes

(3) Remission: number of patients who remit on treatment, based on the endpoint absolute status of the patients, as measured by MADRS, HAM-D, or any other validated depression scale. Examples of definitions of remission include 7 or less on 17-item HAM-D (Furukawa et al. , 2007, Tohen et al. , 2009). The definitions of remission on MADRS differ according to the primary diagnosis. It is defined as 11 or less on the score in major depressive disorder (Bandelow et al. , 2006) and as 7 or less in bipolar disorder (Tohen, Frank, 2009); we will accept the study authors' original definition. If the original authors report several outcomes corresponding with our definition of remission, we will give preference to MADRS.
(4) Severity of depression symptoms, based on a continuous outcome of group mean scores at the end of treatment using MADRS, HAM-D, or any other validated depression scale.
(5) The number of participants who switched their mood episode to mania/hypomania association with the treatment.
(6) The number of participants who completed suicide or made a suicide attempt.
(7) Frequency, severity and category of other clinically significant adverse events.

Search methods for identification of studies

Electronic searches

We will search MEDLINE, PsycINFO, EMBASE and the Cochrane Central Register of Controlled Trials (CENTRAL).
In order to identify randomized trials, we will use the search term of the cochrane highly sensitive search strategy for identifying randomized trials in each databases with sensitivity-maximizing version (Cochrane handbook).
Together with RCT filters, we will search generic terms for bipolar depression and individual drug names; Carbamazepine, Clonazepam, Eslicarbazepine, Felbamate , Gabapentin, Lamotrigine, Levetiracetam, Licarbazepine, Lithium, Oxcarbazepine, Phenytoin, Pregabalin, Retigabine/Ezogabine, Tiagabine, Topiramate, Valproate/divalproex/valproic acid, Vigabatrine, Zonisamide, Amisulpride, Aripiprazole, Asenapine,

Bifeprunox, Blonanserin, Cariprazine, Chlorpromazine, Clozapine, Fluphent(h)ixol, Haloperidol, Loxapine, Lurasidone, Olanzapine/OFC, Paliperidone, Perospirone, Perphenazine, Pimozide, Quetiapine, Risperidone, Sulpiride, Sertindole, Ziprasidone, Zotepine, Sultopride, Paroxetine, Fluoxetine, Fluvoxamine, Sertraline, Citalopram, Escitalopram, Venlafaxine, Desvenlafaxine, Duloxetine, Milnacipran, Bupropion, Mirtazapine, Reboxetine, Amitriptyline, Amoxapine, Imipramine, Desipramine, Nortriptyline, Clomipramine, Trimipramine, Doxepin, Protryptyline, Maprotiline, Butriptyline, Dosulepin, Lofepramine, Mianserin, Oxaprotiline, Setiptiline, Moclobemide, Phenelzine, Selegiline, Rasaligine, Tranylcypromine, Isocarboxazid, Brofaromine, Agomelatine, Tianeptine, Buspirone, Gepirone, Etoperidone, Nisoxetine, Trazodne, Viloxazine, Zimelidine, Vilazodone, Modafinil, Methylphenidate, Pemoline, Pramipexole, Tamoxifen, Ketamine, Minocycline, Lisdexamfetamine, Riluzole, Memantine, Amantadine, Verapamil, Liothyronine(T3), Lu AA39959, Lu AA34893, Ceftriaxone.

References to trials are also sourced from international trials registers via the World Health Organization's trials portal (http://apps.who.int/trialsearch/); regulatory agencies; drug companies; the hand-searching of key journals, conference proceedings and other (non-Cochrane) systematic reviews and meta-analyses. No language restriction will be applied.

Searching other resources

In addition to electronic searches, we will contact other resources to obtain unpublished trail evidences (Chan, 2012).

1. Reference lists

The references of all selected studies will be searched for more published reports and citations of unpublished studies. Relevant review papers will be checked.

2. Personal communication

Subject experts will be contacted to check that all relevant studies, either published or unpublished, have been considered for inclusion.

Data collection

Selection of studies

Two review authors will examine the abstracts of all publications obtained through the search strategy. Full articles of all the studies identified by either of the review authors will then be obtained and inspected by the same two review authors to identify trials meeting the following criteria:

1. Double-blind randomized controlled trial;
2. Participants with major depressive episode in bipolar I, or II disorder diagnosed by operationalized criteria; and
3. Pharmacological intervention with prescription drugs

Conflicts of opinion regarding eligibility of a study will be discussed with a third review author, having retrieved the full paper and consulted the authors if necessary, until consensus is reached. External subject or methodological experts will be consulted if necessary.

Data extraction and management

Data from each study will be extracted independently by two review authors. Degree of agreement between the two independent raters will be reported in terms of kappa coefficients and percentage agreements for main outcomes and risk of bias items. Any disagreement will be resolved through discussion and in consultation with the principal investigators. Where necessary, the authors of the studies will be contacted for further information.

Information relating to study population, sample size, interventions, comparators, potential biases in the conduct of the trial, and outcomes will be abstracted from the original reports into specially designed paper forms then entered into a spreadsheet.

Assessment of risk of bias in included studies

Risk of bias will be assessed for each included study using the Cochrane Collaboration 'risk of bias' tool (Higgins and Green, 2011). The following 7 domains will be considered:

1. Sequence generation: was the allocation sequence adequately generated?

2. Allocation concealment: was allocation adequately concealed?
3. Blinding of participants, personnel and outcome assessors for each main outcome or class of outcomes: was knowledge of the allocated treatment adequately prevented during the study?
4. Incomplete outcome data for each main outcome or class of outcomes: were incomplete outcome data adequately addressed?
5. Selective outcome reporting: are reports of the study free of suggestion of selective outcome reporting?

We also assessed

6. Sponsorship bias.
7. Other sources of bias included but are not limited to:
   - Suboptimal randomization, such as recruiting additional patients to one arm which had a large number of dropouts
   - Stopped early due to some data-dependent process (including a formal-stopping rule)
   - Had extreme baseline imbalance
   - Differential treatment duration among the arms
   - Insufficient delivery of treatment or insensitive scales to measure outcomes, leading to null results

A description of what was reported to have happened in each study will be provided, and a judgment on the risk of bias will be made for each domain, based on the following three categories:

- High risk of bias
- Low risk of bias
- Unclear risk of bias.

Two independent review authors will assess the risk of bias in selected studies. Degree of agreement between the two independent raters will be reported. Any disagreement will be resolved through discussion and in consultation with the principal investigators. Where necessary, the authors of the studies will be contacted for further information.

Statistical analyses

Measures of treatment effect

Considering that depression trials are usually small and that data distribution is difficult to assess for studies with small samples, in this review priority will be given to the use and analysis of dichotomous variables both for efficacy and acceptability.

Dichotomous outcomes: these outcomes will be analyzed by calculating a pooled odds ratio (OR) and 95% confidence intervals for each comparison.

Continuous outcomes: Where different measures are used to assess the same outcome, data will be pooled with standardized mean difference (SMD) and 95% confidence intervals calculated.

Dealing with missing data

Missing dichotomous data will be managed according to the intention to treat (ITT) principle, and it will be assumed that patients in the full analysis set who dropped out after randomization had a negative outcome. When dichotomous efficacy outcomes are not reported but baseline mean and endpoint mean and standard deviation of the depression rating scales are provided, we will calculate the number of responding patients at 8 weeks (range 4 to 12 weeks) employing a validated imputation method (Furukawa et al. , 2005). We are aware that other methods to impute response rate are available and have been investigated (Anzures-Cabrera et al. , 2011). Even though these imputation methods are valid and may give odds ratios (ORs) with narrower CIs, they only produce logORs and their variances rather than raw data. As we opt for a model based on 2x2 tables using the binomial likelihood, the Furukawa method will be used in our review.

Missing continuous data will either be analyzed on an endpoint basis, including only participants with a final assessment, or analyzed using last observation carried forward to the final assessment (LOCF) if LOCF data were reported by the trial authors. When P values, t-values, confidence intervals or standard errors are reported in articles, SD will be calculated from their values (Altman and Bland, 1996). Where SDs are missing, attempts will be made to obtain these data through contacting trial authors. Where SDs are not available from trial authors and the vast majority of actual SDs are available and only a minority of SDs are unavailable or unobtainable, a method used for imputing SDs and calculating percentage responders devised by Furukawa and colleagues (Furukawa et al. , 2006) will be used. We will check that the original standard deviations are properly distributed, so that the imputed standard deviation represents the average. Where this

method is employed, data will be interpreted with caution, taking account of the degree of heterogeneity observed. A sensitivity analysis will also be undertaken to examine the effect of the decision to use imputed data.
Where additional figures are not available or obtainable, and it is not deemed appropriate to use the Furukawa method described above, the study data will not be included in the comparison of interest.

Data synthesis

We will generate descriptive statistics for trial and study population characteristics across all eligible trials, describing the types of comparisons and some important variables, either clinical or methodological (such as year of publication, age, severity of illness, sponsorship, clinical setting).

Pair-wise meta-analysis

For each pair-wise comparison between treatments, the odds ratio will be calculated with a 95% CI. A standard, pair-wise meta-analysis will be conducted for each pair-wise comparison of treatments. We plan to use a random-effects model to incorporate the assumption that the different studies are estimating different, yet related, treatment effects (DerSimonian and Laird, 1986). Where there are <3 studies we will combine in a fixed effect analysis (Borenstein et al. , 2009, Mantel and Haenszel, 1959).
A prediction interval, which captures the uncertainty in the summary estimate, the estimate of the between study standard deviation (Tau) and the uncertainty in Tau (Higgins et al. , 2009), will also be estimated.

MTM

To ensure that the network is connected, a network diagram will be constructed for all the outcomes. Note that MTM is only possible for a connected set of treatments.
Random-effects MTM, taking into account the heterogeneity of treatment effects across studies will be conducted in a Bayesian framework using Markov Chain Monte Carlo methods in OpenBUGS 3.2.1 (http://www.openbugs.info/w/FrontPage). MTM combines direct and indirect evidence for any given pair of treatments, and takes into account correlation induced by multi-arm trial. Results for the comparative efficacy and acceptability are presented by OR estimates and 95% confidence intervals (approximately computed by posterior means and 95% probability intervals). We also evaluate the ranking of efficacy and tolerability using posterior probability which treatment is the most efficacious regimen, the second best, the third best, and so on. The goodness of fit of the model to the data will be measured by the posterior mean of the residual deviance. This is defined as the difference between the deviance for the fitted model and the deviance for the saturated model, where deviance measures the fit of the model to the data points using the likelihood function. We will examine leverage plots to help identify any specific data points (trial arms) that were fitting poorly in each model. A leverage plot displays the leverage (a measure of influence equal to the contribution of each trial arm to P_D, the effective number of parameters) versus the signed, square root of the residual deviance (a measure of fit) for each data point. Points with a high leverage are influential, which means that they have a strong influence on the model parameters that generate their fitted values.

Unit of analysis issues

We will correct for intra-class correlation of cluster-randomized trials if sufficient information is available, and otherwise, we will also conduct MTM for the summary measures (logOR or SMD), directly.

Assessment of heterogeneity

Pair-wise meta-analyses

Visual inspection of the forest plots will be used to investigate the possibility of statistical heterogeneity. This will be supplemented using the I-squared statistic. This provides an estimate of the percentage of variability due to heterogeneity rather than a sampling error (Higgins et al. , 2003). 95% confidence intervals will be calculated for I-squared, and a P value from a χ^2 test for heterogeneity will be used to assess evidence of its presence. We will also report Tau^2 (the between study variance). We consider a degree of heterogeneity inevitable and therefore only I^2 values $\geq$50% will be explored further using subgroup analyses for the primary outcome only.

MTM

Inconsistency can be considered an additional layer of heterogeneity which can occur in networks of evidence. It can occur when there is a discrepancy between a direct and indirect estimate of treatment effect. As such inconsistency is considered a property of 'closed loops' of evidence. As a first step, we will calculate the difference between indirect and direct estimates in each closed loop formed by the network of trials as a measure of inconsistency and we will subsequently examine whether there are any material discrepancies. We will also use model fit statistics as an informal check of inconsistency. In case of considerable inconsistency we will investigate possible sources of it. (See below for possible sources).
In the network meta-analysis we will assume homogeneous between study variability across studies (Higgins and Whitehead, 1996). We will report Tau (the standard deviation of underlying effects across studies) as our estimate of heterogeneity. We will also report the effective number of parameters, pD , which increases with the degree of heterogeneity in the random effect models, and so can also be viewed as a measure of heterogeneity.

Assessment of reporting biases

In order to minimize the impact of reporting bias, we will undertake comprehensive search of studies from multiple sources (including trial registries), increasing efforts to identify unpublished material including contacts with the original study authors, and including non-English language publications.
We will consider using funnel plots to assess the impact of reporting bias on the estimates of treatment effect with visual inspections. We will use Tanaka's method (manuscript in preparation) and Chaimani's method (Chaimani and Salanti, 2012) to assess small study effects in MTM.

Sources of inconsistency and heterogeneity

The following sources of possible clinical heterogeneity are listed a priori and will be examined as effect modifiers in the MTM in meta-regression.

1. Year of publication
2. Baseline depression severity: MADRS or 17-item HAMD scores at baseline will be entered as a covariate.
3. %Female
4. Co-intervention allowed
5. Sponsorship
6. %Participants who assign to placebo
7. Break-through episode
8. Mean age at onset

Subgroup analyses

We will conduct subgroup analyses with the following variables.

1. Bipolar disorder subtype

The efficacy and tolerability to pharmacological treatment will be different between bipolar I and II disorder.

2. Rapid-cycling bipolar disorder

Rapid-cycling bipolar disorder will be specified when the patient with bipolar disorder experienced more than four episodes per year. It appears to have less response to pharmacological treatment.

Sensitivity analyses

In order to assess sensitivity of the random-effects model, the fixed effect model analyses will also be conducted.
In order to examine if the obtained results are preserved when we limit the included studies to high quality ones only, we will examine the following variables.

1. Allocation concealment

Allocation concealment will be used as a marker of trial quality (Wood et al. , 2008). Studies that were rated at high or unclear risk of bias for allocation concealment will be excluded.

2. Imputation

Trials where missing data have been imputed will be excluded.

References

Altman DG, Bland JM. Detecting skewness from summary information. Bmj. 1996;313:1200.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th Edition. Washington, D.C.: American Psychiatric Association; 1994.

Angst J. [On the etiology and nosology of endogenous depressive psychoses. A genetic, sociologic and clinical study]. Monographien aus dem Gesamtgebiete der Neurologie und Psychiatrie. 1966;112:1-118.

Anzures-Cabrera J, Sarpatwari A, Higgins JP. Expressing findings from meta-analyses of continuous outcomes in terms of risks. Statistics in medicine. 2011;30:2967-85.

Bandelow B, Baldwin DS, Dolberg OT, Andersen HF, Stein DJ. What is the threshold for symptomatic response and remission for major depressive disorder, panic disorder, social anxiety disorder, and generalized anxiety disorder? J Clin Psychiatry. 2006;67:1428-34.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis: John Wiley & Sons; 2009.

Cade JF. Lithium salts in the treatment of psychotic excitement. The Medical journal of Australia. 1949;2:349-52.

Calabrese JR, Hirschfeld RM, Frye MA, Reed ML. Impact of depressive symptoms compared with manic symptoms in bipolar disorder: results of a U.S. community-based sample. J Clin Psychiatry. 2004;65:1499-504.

Chaimani A, Salanti G. Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. Research Synthesis Methods. 2012.

Chan AW. Out of sight but not out of mind: how to search for unpublished clinical trial evidence. BMJ. 2012;344:d8013.

Chiesa A, Chierzi F, De Ronchi D, Serretti A. Quetiapine for bipolar depression: a systematic review and meta-analysis. Int Clin Psychopharmacol. 2012;27:76-90.

Cipriani A, Barbui C, Salanti G, Rendell J, Brown R, Stockton S, et al. Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments meta-analysis. Lancet. 2011.

Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet. 2009;373:746-58.

Dean BB, Gerner D, Gerner RH. A systematic review evaluating health-related quality of life, work impairment, and healthcare costs and utilization in bipolar disorder. Current medical research and opinion. 2004;20:139-54.

DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled clinical trials. 1986;7:177-88.

Furukawa TA, Akechi T, Azuma H, Okuyama T, Higuchi T. Evidence-based guidelines for interpretation of the Hamilton Rating Scale for Depression. J Clin Psychopharmacol. 2007;27:531-4.

Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. J Clin Epidemiol. 2006;59:7-10.

Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analyses. Int Clin Psychopharmacol. 2005;20:49-52.

Geddes JR, Calabrese JR, Goodwin GM. Lamotrigine for treatment of bipolar depression: independent meta-analysis and meta-regression of individual patient data from five randomised trials. Br J Psychiatry. 2009;194:4-9.

Goodwin FK, Jamison KR. Manic-Depressive Illness. 2nd ed: Oxford University Press; 2007.

Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated September 2011] Available from www.cochrane-handbook.org. 2011.

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Bmj. 2003;327:557-60.

Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc. 2009;172:137-59.

Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. Statistics in medicine. 1996;15:2733-49.

Judd LL, Akiskal HS, Schettler PJ, Coryell W, Endicott J, Maser JD, et al. A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar II disorder. Arch Gen Psychiatry.

2003;60:261-9.
Judd LL, Akiskal HS, Schettler PJ, Endicott J, Maser J, Solomon DA, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. Arch Gen Psychiatry. 2002;59:530-7.
Keller MB, Lavori PW, Coryell W, Endicott J, Mueller TI. Bipolar I: a five-year prospective follow-up. The Journal of nervous and mental disease. 1993;181:238-45.
Leonhard K. The Classification of Endogenous Psychoses. 5th ed. New York, Irvington1979.
Lumley T. Network meta-analysis for indirect treatment comparisons. Statistics in medicine. 2002;21:2313-24.
Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute. 1959;22:719-48.
Perris C. A study of bipolar (manic-depressive) and unipolar recurrent depressive psychoses. Acta psychiatrica Scandinavica Supplementum. 1966;194:9-189.
Psaty BM, Lumley T, Furberg CD, Schellenbaum G, Pahor M, Alderman MH, et al. Health outcomes associated with various antihypertensive therapies used as first-line agents: a network meta-analysis. Jama. 2003;289:2534-44.
Sachs GS, Nierenberg AA, Calabrese JR, Marangell LB, Wisniewski SR, Gyulai L, et al. Effectiveness of adjunctive antidepressant treatment for bipolar depression. The New England journal of medicine. 2007;356:1711-22.
Sidor MM, Macqueen GM. Antidepressants for the acute treatment of bipolar depression: a systematic review and meta-analysis. J Clin Psychiatry. 2011;72:156-67.
Stettler C, Wandel S, Allemann S, Kastrati A, Morice MC, Schomig A, et al. Outcomes associated with drug-eluting and bare-metal stents: a collaborative network meta-analysis. Lancet. 2007;370:937-48.
Suppes T, Leverich GS, Keck PE, Nolen WA, Denicoff KD, Altshuler LL, et al. The Stanley Foundation Bipolar Treatment Outcome Network. II. Demographics and illness characteristics of the first 261 patients. J Affect Disord. 2001;67:45-59.
Tohen M, Frank E, Bowden CL, Colom F, Ghaemi SN, Yatham LN, et al. The International Society for Bipolar Disorders (ISBD) Task Force report on the nomenclature of course and outcome in bipolar disorders. Bipolar disorders. 2009;11:453-73.
Tohen M, Waternaux CM, Tsuang MT. Outcome in Mania. A 4-year prospective follow-up of 75 patients utilizing survival analysis. Arch Gen Psychiatry. 1990;47:1106-11.
Trede K, Salvatore P, Baethge C, Gerhard A, Maggini C, Baldessarini RJ. Manic-depressive illness: evolution in Kraepelin's Textbook, 1883-1926. Harvard review of psychiatry. 2005;13:155-78.
Vieta E, Locklear J, Gunther O, Ekman M, Miltenburger C, Chatterton ML, et al. Treatment options for bipolar depression: a systematic review of randomized, controlled trials. J Clin Psychopharmacol. 2010;30:579-90.
WHO. ICD-10 Classifications of Mental and Behavioural Disorder: Clinical Descriptions and Diagnostic Guidelines. Geneva1993.
WHO. The global burden of disease: 2004 update. http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/index.html2008.
Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. Bmj. 2008;336:601-5.