COMPARATIVE EFFICACY AND TOLERABILITY OF PHARMACOLOGICAL TREATMENTS IN THE MAINTENANCE TREATMENT OF BIPOLAR DISORDER: A MULTIPLE TREATMENTS 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

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).

Description of the intervention

Bipolar disorder is a chronic relapsing and remitting disorder and long-term treatment is often required to minimise the risk of recurrence. Many clinical guidelines recommended continuing maintenance treatments, after successful treatment for acute phase of a mood episode in bipolar disorder [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.]. Although some specific psychological treatments, including cognitive behavioral therapy (CBT) (Otto et al. , 2003), family-focused therapy (Miklowitz, 2008), interpersonal social rhythm therapy (IPSRT) (Frank, 2005) and psychoeducation (Colom et al. , 2009) have been shown some effect for preventing relapse, all of them are provided in conjunction with some drugs and pharmacological interventions remain an essential part of maintenance treatment in bipolar disorder.

Lithium and some anti-epileptics (valproic acid, carbamazepine, oxcarbazepine and lamotrigine) are typical agents used for long term treatment of bipolar disorder. While anti-psychotics have been typically used for the treatment of acute mania, recent studies have demonstrated that some of them, especially second-generation antipsychotics, also were efficacious in maintenance treatment. Antipsychotics more likely to have prophylactic effect on relapse of manic or mixed episodes rather than on depressive ones, but quetiapine, one of the second-generation antipsychotics have been reported to prevent relapse of both mood episodes (Weisler et al. , 2011). By contrast, antidepressants have been used for the treatment of acute bipolar depression, and, notwithstanding the risk of switching to high mood episodes, there is some evidence that some antidepressants are effective in preventing depressive relapse in bipolar disorder.

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 maintenance treatment in bipolar disorder. 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. Some reported the result as a pooled risk ratio, and another reported it as a pooled hazard ratio with conventional methods of meta-analysis, that shows a direct head-to-head comparison in effect size between two interventions.

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 a relative efficacy of the two treatments can be estimated even when there is no study that directly compare the efficacy between two interventions.

In conducting MTM we need to pay particular attention to transitivity/consistency of the network and comparability of included patients across nodes of the network. In the case of bipolar maintenance studies, there is especial risk of heterogeneity because studies would include different mix of bipolar I and II patients, some would deal with maintenance after manic episodes and others after depressive episodes, some would allow psychosocial co-intervention and others focus on drug therapies only and so on. In order to be as broadly representative of bipolar maintenance treatments as possible, we will include all the relevant studies recognizing that differences in patient and intervention characteristics might introduce heterogeneity and we will employ random effects to reflect differences across studies. If the distributions of characteristics are comparable across comparisons of interventions, as we expect them to be, then transitivity is likely to hold. We will evaluate the assumption of transitivity by statistical means and we will employ network meta-regression to account for small differences in effect modifiers across comparisons and studies.
One MTM study in the treatment for the maintenance phase in bipolar disorder has been reported in the literature (Soares-Weiser et al. , 2007), but many important evidences in this field have emerged since then, especially concerning the treatment effect of second-generation anti-psychotics. We are therefore in dire need of a comparative effectiveness research with MTM including such newer pharmacological treatment in the maintenance treatment of bipolar disorder to inform our clinical practices.

Objectives

To compare the efficacy and tolerability of different drugs or combinations of drugs in the maintenance treatment of bipolar disorder in adults.

Methods

Criteria for considering studies for this review

Types of studies

Inclusion criteria
Randomized controlled trials (RCTs) comparing with placebo, or active comparator with at least 3 months of follow-up in the maintenance phase treatment of 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.
We will include open RCTs but will run a sensitivity analysis to examine the effect of our decision by limiting the analyses to double-blind RCTs.
Exclusion criteria
Quasi-randomized controlled trials, in which treatment assignment is decided through methods such as alternate days of the week, will be excluded. Studies in which participants are randomized to maintenance treatment while in the acute phase (the so-called, continuation studies) will be excluded.

Types of participants

Participant characteristics

Inclusion criteria
Participants aged 18 or older, of both sexes with a primary diagnosis of 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 outpatient settings will be included. We will also include studies that participants are continued to follow-up in outpatient settings after being randomized during hospitalization.

Types of interventions

Included interventions

We will include all the pharmacological interventions in the maintenance treatment of bipolar I, or II disorder, even if they are not licensed in any countries.

Excluded interventions

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 any mood episode (depressive, manic or mixed relapse) as defined by author at the longest available follow-up.
(2) Treatment tolerability: the number of participants who drop out of the treatment due to adverse events at the longest available follow-up.

Secondary outcomes

(3) The number of participants who completed suicide or made a deliberate self harm.
(4) The number of participants who had serious adverse event
(5) Social functioning

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).
Studies in maintenance treatment for bipolar disorder will be searched using following term in combination with individual drug names: [bipolar disorder OR mania OR manic OR cyclothym* OR hypomani* OR hypo-mani* OR rapid cycl* OR rapid-cycl*) AND (maintenance OR prophylaxis OR prevention OR preventive OR recurrence OR relapse OR long term OR long-term)].

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 evidence (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. Randomized controlled trial;
2. Participants with bipolar I, or II disorder diagnosed by operationalized criteria; and
3. Pharmacological intervention

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.

We will use data at the longest available follow-up.

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

In studies of maintenance phase treatment of bipolar disorder, freedom from mood episodes may be regarded as a time-to-event data. Time-to-event data may be expressed either in the form of hazard ratio, or as cumulative count: the total number of participants who have experienced an event at a specific time point.

In this review priority will be given to the use of hazard ratio data when both statistics are available, because it account for censoring, incorporate time to event information, and may be adjusted for co-variables.
Log hazard ratio will be estimated using methods of Parmer et al. in pair-wise meta-analysis when hazard ratio is not reported (Parmar et al. , 1998). In multiple treatment meta-analysis, hazard ratio and cumulative count data will be combined in a single analysis on the hazard ratio scale (Woods et al. , 2010).
A pooled hazard ratio (HR) and 95% confidence intervals for each comparison will be calculated.

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 hazard 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). 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). Data on survival endpoints are combined on the hazard ratio scale by the methods of Woods et al (2010). 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 tolerability are presented by HR 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.
We plan to use two sets of data for MTM analyses. First, data set that only includes the evidences from studies comparing mono therapy will be used. If the pooled HR of lithium arm and that of VPA arm would be comparable each other and against placebo in first analysis, lithium arm and VPA arm will be combined together to be mood stabilizer arm and then second MTM analysis with all the evidences including combination therapy will be conducted as main analysis.

Unit of analysis issues

We will correct for intra-class correlation if we must include cluster-randomized trials.

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 ≥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. %Female
3. Type of recent mood episode before randomization
4. Medication used in the recent mood episode before randomization
5. Co-intervention allowed
6. Sponsorship
7. Presence of placebo control

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. Blinding

More treatment effects will be observed in the studies with lack of blinding, by 9% on average measured as odds ratio (Pildal et al. , 2007). That will be especially the case with more subjective outcomes (Wood, Egger, 2008). Studies will be limited to those stated double-blinded.

3. Length of follow-up

Some maintenance studies have followed the patients up to 6 months only, which may be regarded as too short for the maintenance treatment of bipolar disorder that often can be life-long. We will exclude studies that have less than 12 months of follow-up.

4. Imputation

Trials where missing data have been imputed will be excluded.

5. Fixed-effects model

In pair-wise comparisons, we will combine the treatment effects of different studies with fixed-effects model.

References

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

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.

Colom F, Vieta E, Sanchez-Moreno J, Palomino-Otiniano R, Reinares M, Goikolea JM, et al. Group psychoeducation for stabilised bipolar disorders: 5-year outcome of a randomised clinical trial. Br J Psychiatry. 2009;194:260-5.

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.

Frank E. Treating Bipolar Disorder: A Clinician's Guide to Interpersonal and Social Rhythm Therapy. New York: The Guilford Press; 2005.

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.

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.

Lumley T. Network meta-analysis for indirect treatment comparisons. Statistics in medicine. 2002;21:2313-24.

Miklowitz DJ. Bipolar Disorder: A Family-Focused Treatment Approach. 2nd ed. New York: The Guilford Press; 2008.

Otto MW, Reilly-Harrington N, Sachs GS. Psychoeducational and cognitive-behavioral strategies in the management of bipolar disorder. J Affect Disord. 2003;73:171-81.

Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Statistics in medicine. 1998;17:2815-34.

Pildal J, Hrobjartsson A, Jorgensen KJ, Hilden J, Altman DG, Gotzsche PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. International journal of epidemiology. 2007;36:847-57.

Soares-Weiser K, Bravo Vergel Y, Beynon S, Dunn G, Barbieri M, Duffy S, et al. A systematic review and economic model of the clinical effectiveness and cost-effectiveness of interventions for preventing relapse in people with bipolar disorder. Health technology assessment. 2007;11:iii-iv, ix-206.

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, 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.

Weisler RH, Nolen WA, Neijber A, Hellqvist A, Paulsson B, Trial 144 Study I. Continuation of quetiapine versus switching to placebo or lithium for maintenance treatment of bipolar I disorder (Trial 144: a randomized controlled study). J Clin Psychiatry. 2011;72:1452-64.

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.

Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. BMC medical research methodology. 2010;10:54.