Editorial

Network meta-analysis: drawing conclusions regarding trials of psychosocial interventions for bipolar disorder

David J. Miklowitz, Andrea Cipriani and Guy M. Goodwin

Summary

Network meta-analysis (NMA) is a statistical technique for making direct and indirect comparisons between different treatment and control groups. Despite its many advantages, NMA may be misleading when evaluating networks that are disconnected, inconsistent or of low reliability and validity. We review how well the analysis of trials of adjunctive psychosocial treatment in bipolar disorder is served by NMA. We conclude with recommendations for future treatment trials in bipolar disorder and guidelines for NMAs.

David J. Miklowitz (pictured) is a professor of psychiatry at the University of California, Los Angeles and specialises in family processes and interventions relevant to bipolar disorder. His work focuses on testing and understanding the mechanisms of adjunctive psychosocial interventions in adults, adolescents and children who have or are at risk for bipolar disorder. Andrea Cipriani is an associate professor at the Department of Psychiatry of the University of Oxford and an honorary consultant psychiatrist at Oxford Health NHS Foundation Trust. His interest is in the methodology of evidence synthesis, with a specific focus on network meta-analysis and individual patient data meta-analysis to inform the mental healthcare decision-making process. Guy M. Goodwin is a senior research fellow and emeritus professor at the Department of Psychiatry of the University of Oxford and an honorary consultant psychiatrist at Oxford Health NHS Foundation Trust. His interest is in quantification of mental states and the treatment of bipolar disorder.

The investigation of effective psychosocial interventions for patients with bipolar disorder has been ongoing since the late 1970s. Various psychotherapy modalities have been manualised and tested as adjuncts to pharmacotherapy in randomised controlled trials (RCTs), with the majority concerning cognitive–behavioural therapy (CBT), interpersonal and social rhythm therapy (IPSRT), family-focused therapy or structured group psychoeducation. There is a consensus that psychoeducation is a key component of good clinical care in bipolar disorder.[1,2] However, more specific conclusions have provoked controversy.

The National Institute for Health and Care Excellence's recommendations, based on multiple meta-analyses, appear to have gone beyond the evidence in an effort to compare the options.[3] The central difficulty is that the trial literature is very heterogeneous, with some studies restricted to euthymic or recovered patients with bipolar I or II disorder and others including patients recruited during an acute manic episode. Many studies use a treatment-as-usual (TAU) control without specifying the features of this comparator. There are no standardised, consensus instruments for assessing change in mood symptoms or functioning. As a result, we know relatively little about what treatments work best for what patients, what outcomes are affected by what treatments, or whether certain treatment structures (for example groups) are more cost-effective than others (for example individual or family). Putative moderators have come and gone, effect sizes that looked impressive in early trials become much smaller on replication, and data concerning treatment mechanisms are lacking. More studies are certainly required to address these uncertainties, but how much is there further to learn from the studies that have already been done?

Network meta-analysis

The answer may lie in network meta-analysis (NMA), a sophisticated statistical technique that aims to synthesise comparative evidence across a network of RCTs (all of which meet quality standards) and rank the different treatment options against each other. The key feature of NMA is that it allows the simultaneous analysis of data comparing treatments within the same study (direct evidence) with data comparing interventions across different studies (indirect evidence). Indirect evidence is important because, using a common comparator such as TAU, it fills the gaps (i.e. treatments that have not been compared directly can be compared indirectly). When combined with direct evidence, indirect evidence increases the precision of the effect estimate between two interventions.[4] NMA is becoming very popular in the scientific literature, but unfortunately such interest also increases the likelihood of adopting poorer methodological quality.[5] As we explain below, NMA results are only valid when the network of comparisons is coherent and well connected.

As for any meta-analysis based on a systematic review, criteria for including and excluding individual studies in these networks must be convergent with the original review question to reduce the risk of spurious findings.[2] Ideally, all studies should have similar designs and, critically, select similar patients. NMAs should also offer formal assessments of the overall quality of the evidence retrieved.

An example of NMA as applied to trials of psychosocial interventions for bipolar disorder

We illustrate these issues concerning network interconnectedness, rating study quality, and inclusion/exclusion decisions in the paper published in the *BJPsych* by Chatterton and colleagues,[6] the first NMA of the bipolar/psychosocial treatment literature. They drew conclusions based on the prespecified outcomes of relapse, depressive and manic symptoms, global functioning and medication adherence (PROSPERO ID: CRD42015016975), each of which, by definition, is associated with a distinct network (i.e. not all studies included in the network contribute to all the outcomes). In the 41 RCTs in bipolar disorder that were included, Chatterton *et al* concluded that (a) caregiver-focused interventions significantly reduced the risk of manic and depressive relapses (risk ratio (RR) = 0.61, 95% CI 0.44 to 0.86); (b) psychoeducation, given alone or in combination with CBT reduced non-adherence with medication (RR = 0.27, 95% CI 0.14 to 0.53 and 0.14, 0.02–0.85 respectively); and (c) the combination of psychoeducation and CBT reduced manic symptoms and increased global functioning (standardised mean difference (SMD) = −0.95, 95% CI −1.47 to −0.43).

None of the interventions were found to be effective in reducing depressive symptoms. This is unexpected: one of the main justifications for developing and using structured psychotherapy in conjunction with medications for bipolar disorder has been that the latter are less effective in controlling depression than mania. Several of the best-studied approaches to bipolar disorder were derived from treatments for patients with unipolar depression. Moreover, the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD), a relatively large-scale multicentre RCT ($n = 293$), directly compared three active, evidence-based psychosocial interventions (CBT, IPSRT and family-focused therapy) with a brief psychoeducational control in adults with bipolar disorder recovering from a major depressive episode.[7] The three intensive therapies were more effective than the brief psychoeducation in decreasing time to depression recovery and increasing likelihood of being well over the study year. However, the STEP-BD trial was excluded from this NMA.

In the Chatterton *et al* analysis, the network of studies that were included is very sparse: there are few direct comparisons of one active intervention to another, so the conclusions are drawn primarily from indirect comparisons between active treatments and TAU. Moreover, few trials adequately specified the TAU condition, which varied from pharmacotherapy only (with no clarification of visit frequency) to pharmacotherapy with case management, to pharmacotherapy with supportive individual therapy, to waitlist controls.

The inclusion/exclusion decisions were at times puzzling. The authors explain that many studies were excluded 'owing to the lack of reporting of extractable data'.[6] As the authors acknowledge, no studies of IPSRT were included, even though there are at least four trials testing its effects against inactive or active control conditions. Apparently, only one of these trials made outcome data available for the NMA, and that study was excluded. The investigators also excluded a study of collaborative care with 306 patients treated at US Veterans Administration facilities,[8] but included a similarly designed study of 441 patients in a health management organisation who received a nearly identical collaborative care treatment.[9] Further, the authors excluded a large Canadian pragmatic trial ($n = 204$) that provided a direct comparison of CBT with group psychoeducation.[10] Large-scale pragmatic trials that compare two or more evidence-based

psychosocial treatments are especially relevant to the direct networks being tested in NMAs.

One limitation of NMA methodology, then, is that the decision rules on which studies to include or exclude may strongly influence the results. Indeed, the eligibility criteria for this NMA seemed to favour smaller studies that test innovations in CBT or group psychoeducational treatment compared with usual care. One is left to wonder how to translate the findings into recommendations for clinical practice given the exclusion of key studies. At minimum, sensitivity analyses should be performed to determine how including or excluding specific studies changes the results.

In this NMA, risk of bias was examined using the Cochrane Collaboration tool and a total quality score was calculated as the sum of all components (a value between 0 and 1, calculated for each study by dividing the total quality score by the highest scoring study in the group). Rating the risk of bias of individual studies is important but not enough. Guidance exists on how to properly rate the quality of evidence supporting treatment effect estimates obtained from NMAs.[11] GRADE (or another reliable rating system) should become mandatory for NMAs published in high-quality, peer-reviewed journals.[12,13]

What of the conclusions that are reached? The most contentious was: if the primary goal is the prevention of relapse in bipolar disorder, educating and supporting caregivers is the most effective intervention. This conclusion suggests that the patient with bipolar disorder need not be involved in psychosocial sessions in order for treatment to be effective. When unpacking this finding, we learn that there were only five studies of caregiver-only treatments, and of those, only two were included in the network for relapse-as-outcome. One of the five studies was excluded because the carer-focused intervention was compared with an attention control treatment. One of the two included caregiver-focused trials tested the effects of one 2 h educational session for caregivers compared with TAU in reducing relapse rates.[14] It is not clear what other treatments the patients received during this trial. The only other study that met the inclusion criteria was a trial that found lower rates of manic (not depressive) relapse in patients whose caregivers received psychoeducation in groups.[15] Finally – and paradoxically – caregiver-focused interventions were ranked as the most effective in preventing relapse and the least effective in reducing non-adherence to medications; yet preventing manic episodes usually requires consistent adherence to medication. Thus, we remain unconvinced by inferences that are based on a limited number of studies tested in a sparse network.

Guidelines for publishing NMAs

NMA has the potential to draw treatment recommendations from existing data, rather than require more RCTs to answer direct questions about specific contrasts. However, drawing conclusions in NMAs based on non-connected networks is challenging. Careful attention should be paid to the consistency between direct and indirect evidence, when available. Thus, if treatment A beats treatment B and B beats C, the loop is consistent only if A beats C. If, instead, C beats A, the loop is inconsistent. In the study protocol, all NMAs should describe a clear strategy for identifying and addressing such inconsistencies.[16] Inconsistency between direct and indirect sources of evidence should be statistically assessed both globally (by comparison of the fit and parsimony of consistency and inconsistency models) and locally (by calculation of the difference between direct and indirect estimates in all closed loops in the network). Even when the global test

of consistency is acceptable, there may be important local inconsistencies that must be appraised in order to evaluate how the combination of direct and indirect evidence – the so-called mixed estimate – is affected. The node splitting method should also be used to calculate the inconsistency of the model, which separates evidence on a particular comparison into direct and indirect evidence.[16]

Conclusions

We join Chatterton *et al* in emphasising several directions for future research but also caution against the misuse of NMA. First, we need to develop more effective approaches to bipolar depressive symptoms, which remain the primary residual and enduring burden of illness between acute episodes. Second, we need a common assessment battery for studies in bipolar disorder (analogous to the MATRIX battery in schizophrenia, but learning from its limitations) so that treatment effects can be more directly compared in NMAs. Third, trial investigators need to anticipate that individual-level data will be requested to develop or update meta-analyses.[17]

The end goal of meta-analysis is to provide evidence to support treatment decision-making relevant to individual patients with bipolar disorder in a variety of clinical states.[17] In evaluating results, we must keep in mind the individual variables that may affect response to various psychosocial and pharmacological treatments (for example a history of non-adherence), and the matching between the clinician's setting or type of practice and the menu of reasonable options now available for treating patients with bipolar disorder.

David J. Miklowitz, PhD, Department of Psychiatry, University of California, Los Angeles (UCLA) School of Medicine, Los Angeles, California, USA and University Department of Psychiatry, University of Oxford, Oxford, UK; **Andrea Cipriani**, MD, PhD, **Guy M. Goodwin**, FMedSci, University Department of Psychiatry, University of Oxford and Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK

Correspondence: David J. Miklowitz, Department of Psychiatry, UCLA Semel Institute, 760 Westwood Plaza Rm A8-256, Los Angeles, CA 90095, USA. Email: dmiklowitz@mednet.ucla.edu

First received 11 Apr 2017, final revision 9 Jun 2017, accepted 27 Jun 2017

References

1 National Institute for Health and Care Excellence. *Bipolar Disorder: The Assessment and Management of Bipolar Disorder in Adults, Children and Young People in Primary and Secondary Care, Updated Edition*. National Clinical Guideline Number 185. NICE, 2014.

2 Goodwin GM, Haddad PM, Ferrier IN, Aronson JK, Barnes T, Cipriani A, et al. Evidence-based guidelines for treating bipolar disorder: revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacology* 2016; **30**: 495–553.

3 Jauhar S, McKenna PJ, Laws KR. NICE guidance on psychological treatments for bipolar disorder: searching for the evidence. *Lancet Psychiatry* 2016; **3**: 386–8.

4 Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013; **159**: 130–7.

5 Zarin W, Veroniki AA, Nincic V, Vafaei A, Reynen E, Motiwala SS, et al. Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. *BMC Med* 2017; **15**: 3.

6 Chatterton ML, Stockings E, Berk M, Barendregt JJ, Carter R, Mihalopoulos C. Psychosocial therapies for the adjunctive treatment of bipolar disorder in adults: network meta-analysis. *Br J Psychiatry* 2017; **210**: 333–41.

7 Miklowitz DJ, Otto MW, Frank E, Reilly-Harrington NA, Wisniewski SR, Kogan JN, et al. Psychosocial treatments for bipolar depression: a 1-year randomized trial from the Systematic Treatment Enhancement Program. *Arch Gen Psychiatry* 2007; **64**: 419–27.

8 Bauer MS, McBride L, Williford WO, Glick H, Kinosian B, Altshuler L, et al. Collaborative care for bipolar disorder: Part II. Impact on clinical outcome, function, and costs. *Psychiatr Serv* 2006; **57**: 937–45.

9 Simon GE, Ludman EJ, Bauer MS, Unutzer J, Operskalski B. Long-term effectiveness and cost of a systematic care program for bipolar disorder. *Arch Gen Psychiatry* 2006; **63**: 500–8.

10 Parikh SV, Zaretsky A, Beaulieu S, Yatham LN, Young LT, Patelis-Siotis I, et al. A randomized controlled trial of psychoeducation or cognitive-behavioral therapy in bipolar disorder: a Canadian Network for Mood and Anxiety treatments (CANMAT) study. *J Clin Psychiatry* 2012; **73**: 803–10.

11 Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015; **162**: 777–84.

12 Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 2014; **349**: g5630.

13 Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014; **9**: e99682.

14 Bordbar M, Soltanifar A, Talaei A. A short-term family focused psychoeducational program for bipolar disorder in Mashhad. *Iran J Med Sci* 2009; **34**: 6.

15 Reinares M, Sánchez-Moreno J, Fountoulakis KN. Psychosocial interventions in bipolar disorder: what, for whom, and when. *J Affect Disord* 2014; **156**: 46–55.

16 Mavridis D, Giannatsi M, Cipriani A, Salanti G. A primer on network meta-analysis with emphasis on mental health. *Evid Based Ment Health* 2015; **18**: 40–6.

17 Ioannidis JP. Meta-analyses can be credible and useful: a new standard. *JAMA Psychiatry* 2017; **74**: 311–2.