TECHNICAL COMMENTARY

NeuRA Discover. Conquer. Cure.

BIPOLAR DISORDERS LIBRARY

Treatments for bipolar versus unipolar depression

Introduction

Bipolar disorders are a group of disorders characterised by episodes of depression and mania or hypomania. Bipolar disorders described in the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, version 5) include bipolar I disorder involving severe depression and mania, bipolar II disorder involving depression and less severe mania (hypomania), and cyclothymic disorder involving many mood swings, with hypomania and depressive symptoms occurring often and fairly constantly.

Major depressive disorder characterised in the DSM-5 involves five (or more) of the following symptoms to be present and represent a change from previous functioning. At least one of the symptoms must be either depressed mood or loss of interest or pleasure, with no history of mania.

- Depressed mood most of the day, nearly every day
- Diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day
- A change of more than 5% of body weight in a month without intent or decrease or increase in appetite
- Insomnia or hypersomnia
- Fatigue or loss of energy
- Psychomotor agitation or retardation that is observable by others
- Feelings of worthlessness or excessive or inappropriate guilt
- Diminished ability to think or concentrate, or indecisiveness
- Recurrent thoughts of death or a suicide attempt or plan

This topic compares the effectiveness of pharmaceutical treatments for bipolar depression versus major depressive disorder.

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2010 that report results for people with a diagnosis of bipolar or related disorders. Due to the high volume of systematic reviews we have now limited inclusion to systematic meta-analyses. Where no systematic meta-analysis exists for a topic, systematic reviews without meta-analysis are included for that topic. Reviews were identified by searching the databases MEDLINE, EMBASE, and PsycINFO. Hand searching reference lists of identified reviews was also conducted. When multiple copies of review topics were found, the most recent and/or comprehensive review was included.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[1]. Reviews with less than 50% of items checked have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large, there is a dose dependent response or if results are reasonably consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[2]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found two systematic reviews that met our inclusion criteria[3, 4].

- Moderate to low quality evidence suggests small, but clinically significant, effects of improved depressive symptoms with antidepressants for major depressive disorder and paroxetine, aripiprazole, lurasidone, olanzapine, quetiapine, ziprazidone, lithium, lamotrigine or divalproex for bipolar depression.
- Moderate quality evidence suggests no differences in depression severity between people with bipolar or unipolar depression after treatment with antidepressants.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Papakostas GI, Martinson MA, Fava M, Iovieno N

Demographic variables, design characteristics, and effect sizes of randomized, placebo-controlled, monotherapy trials of major depressive disorder and bipolar depression

Journal of Clinical Psychiatry 2016; 77: e619-e24

View review abstract online

Comparison	**All monotherapies for depression in people with acute major depressive disorder or bipolar depression vs. placebo.** **Monotherapies for depression were all antidepressants. Monotherapies for bipolar depression included the antidepressant paroxetine, antipsychotics aripiprazole, lurasidone, olanzapine, quetiapine or ziprazidone, and mood stabilisers/anticonvulsants lithium, lamotrigine or divalproex.** **Mean duration of treatment was ~7 weeks. The completion rate for bipolar disorder was less than for major depressive disorder (~64% vs. 72%).**
Summary of evidence	**Moderate to low quality evidence (inconsistent, some indirectness, precise, large samples) suggests small, but clinically significant, effects of improved depressive symptoms with antidepressants for major depressive disorder and paroxetine, aripiprazole, lurasidone, olanzapine, quetiapine, ziprazidone, lithium, lamotrigine or divalproex for bipolar depression.**

Depression

Monotherapy with antidepressants was significantly more effective than placebo for major depressive disorder (small effect);

196 RCTs, N = 56,133, response rates for antidepressants = 52.7% vs. placebo = 37.5%

RR = 1.373, 95%CI 1.351 to 1.396, $p < 0.001$, Q = 693.827, $p < 0.001$

All monotherapies (mixed drug classes) was significantly more effective than placebo for bipolar depression (small effect);

19 RCTs, N = 7,191, response rates for all drug classes = 54.7% vs. placebo = 40.5%%

RR = 1.257, 95%CI 1.185 to 1.335, $p < 0.001$, Q = 45.263, $p = 0.010$

Meta-regression analysis suggested the risk ratio for major depressive disorder was significantly larger than the risk ratio for bipolar depression ($p = 0.005$ adjusted for symptom severity at baseline and probability of being randomised to placebo). However, the number needed to treat was

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

approximately 7 in both major and bipolar depression trials, which is considered clinically significant for both disorders.	
Consistency‡	Inconsistent
Precision§	Precise
Directness‖	Direct comparison for major depressive disorder vs. placebo (all antidepressants); indirect comparison for bipolar depression vs. placebo (mixed drug classes).

Vazquez G, Tondo L, Baldessarini RJ

Comparison of antidepressant responses in patients with bipolar vs. unipolar depression: a meta-analytic review

Pharmacopsychiatry 2011; 44: 21-6

View review abstract online

Comparison	**Efficacy of antidepressants in people with bipolar vs. unipolar depression.**
Summary of evidence	**Moderate to low quality evidence (mixed study designs, inconsistent, precise, direct, large sample) suggests no differences in depression severity between people with bipolar or unipolar depression after treatment with antidepressants.**
Depression	
No significant differences between bipolar and unipolar depression; 10 studies (mixed designs), N = 3,089, RR = 1.05, 95%CI 0.96 to 1.15, *p* = 0.34 Meta-regression showed no differences in responses according to bipolar disorder subtype, number of subjects per study, sex, age, or length of treatment. Risk of switching to mania averaged 2.50% per week for bipolar disorder patients, with 70% also treated with mood stabilisers, vs. 0.275% per week for unipolar depression patients.	
Consistency	Forest plot appears inconsistent.
Precision	Precise
Directness	Direct

Explanation of acronyms

CI = confidence interval, DSM = American Psychiatric Association's Diagnostic and Statistical Manual, N = number of participants, *p* = probability of rejecting a null hypothesis of no differences between groups, RR = risk ratio, vs. = versus

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[5].

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardsed mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) that allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 and over represents a large effect[5].

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[6]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship between variables. They can provide an indirect indication of prediction, but do not

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[5];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[7].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Papakostas GI, Martinson MA, Fava M, Iovieno N (2016): Demographic variables, design characteristics, and effect sizes of randomized, placebo-controlled, monotherapy trials of major depressive disorder and bipolar depression. *Journal of Clinical Psychiatry* 77: e619-e24.
4. Vazquez G, Tondo L, Baldessarini RJ (2011): Comparison of antidepressant responses in patients with bipolar vs. unipolar depression: a meta-analytic review. *Pharmacopsychiatry* 44: 21-6.
5. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
6. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
7. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au