TECHNICAL COMMENTARY

Cognition in first-episode bipolar disorder

Introduction

Cognitive dysfunction is a common feature of bipolar disorder that exists across a number of cognitive domains and usually persists in remission. It is unclear whether cognitive deficits are apparent prior to the onset of bipolar disorder or whether they develop during the course of the illness. Assessing cognitive ability in people with a first-episode of bipolar disorder helps determine whether cognitive deficits were apparent prior to illness onset.

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 separately for people with a diagnosis of bipolar and related disorders. 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, only the most recent and comprehensive review was included. Reviews with pooled data are prioritised for inclusion.

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 (RCT) 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 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 or if there is a dose dependent response. We have also taken into account sample size and whether results are 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].

- High quality evidence shows medium-sized effects of poorer global cognition and processing speed, and small effects of poorer premorbid IQ, working memory, fluency, and reasoning in people with first-episode bipolar disorder compared to controls.
- Moderate to high quality evidence suggests a large effect of poorer attention, medium-sized effects of poorer current IQ, verbal memory, and visual memory in people with first-episode bipolar disorder compared to controls.
- Compared to people with first-episode schizophrenia, moderate to high quality

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

evidence shows medium-sized effects of better verbal memory, verbal fluency, premorbid IQ, in people with first-episode bipolar disorder. Moderate quality evidence also shows small effects of better working memory and processing speed. Moderate to low quality evidence shows a medium-sized effect of better current IQ with no differences in attention or reasoning.

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

Bora E, Pantelis C

Meta-analysis of Cognitive Impairment in First-Episode Bipolar Disorder: Comparison With First-Episode Schizophrenia and Healthy Controls

Schizophrenia Bulletin 2015; 41(5): 1095-1104

View review abstract online

Comparison 1	**Cognitive functioning in people with first-episode bipolar disorder vs. controls.**
Summary of evidence	**High quality evidence (large samples, direct, precise, consistent) shows medium-sized effects of poorer global cognition and processing speed, and small effects of poorer premorbid IQ, working memory, fluency, and reasoning in people with first-episode bipolar disorder.** **Moderate to high quality evidence (medium-sized samples or inconsistent) suggests a large effect of poorer attention, medium-sized effects of poorer current IQ, verbal memory, and visual memory in people with first-episode bipolar disorder.**

Global cognition

A significant, medium-sized effect of poorer global cognition in people with first-episode bipolar disorder;

15 studies, N = 1,950, d = 0.54, 95%CI 0.41 to 0.66, $p < 0.001$, I^2 = 19.8%, p = 0.23

There were no changes in the effect size according to gender, education, age, state (euthymic vs non-euthymic), and exclusion of people with drug use.

Authors report no evidence of publication bias.

IQ

Significant, small to medium-sized effects of lower premorbid and current IQ in people with first-episode bipolar disorder;

Premorbid IQ: 7 studies, N = 707, d = 0.26, 95%CI 0.10 to 0.42, $p < 0.001$, I^2 = 0%, p = 0.73

Current IQ: 8 studies, N = 890, d = 0.45, 95%CI 0.19 to 0.71, $p < 0.001$, I^2 = 63.6%, p = 0.007

Authors report no evidence of publication bias.

Memory

Significant, small to medium-sized effects of poorer verbal memory, working memory, and visual memory in people with first-episode bipolar disorder; Verbal memory: 6 studies, N = 1,097, *d* = 0.63, 95%CI 0.39 to 0.86, *p* < 0.001, I^2 = 62%, *p* = 0.02 Working memory: 9 studies, N = 10,140, *d* = 0.34, 95%CI 0.20 to 0.47, *p* < 0.001, I^2 = 0.6%, *p* = 0.42 Spatial working memory: 3 studies, N = 271, *d* = 0.38, 95%CI 0.06 to 0.71, *p* = 0.02, I^2 not reported Visual memory: 5 studies, N = 590, *d* = 0.51, 95%CI 0.16 to 0.86, *p* = 0.004, I^2 = 68%, *p* = 0.01 Authors report no evidence of publication bias.	
Processing speed	
A significant, medium-sized effect of poorer processing speed in people with first-episode bipolar disorder; 8 studies, N = 785, *d* = 0.61, 95%CI 0.39 to 0.84, *p* < 0.001, I^2 = 47.2%, *p* = 0.07 Authors report no evidence of publication bias.	
Fluency	
A significant, small effect of poorer fluency in people with first-episode bipolar disorder; 9 studies, N = 1,280, *d* = 0.36, 95%CI 0.17 to 0.55, *p* < 0.001, I^2 = 47%, *p* = 0.06 Authors report no evidence of publication bias.	
Attention	
A significant, large effect of poorer attention in people with first-episode bipolar disorder; 3 studies, N = 262, *d* = 0.80, 95%CI 0.54 to 1.06, *p* < 0.001, I^2 = 0%, *p* = 0.96 Authors report no evidence of publication bias.	
Reasoning	
A significant, small effect of poorer reasoning in people with first-episode bipolar disorder; 7 studies, N = 1,053, *d* = 0.31, 95%CI 0.05 to 0.56, *p* = 0.02, I^2 = 62%, *p* = 0.01 Authors report no evidence of publication bias.	
Consistency in results‡	Consistent, apart from current IQ, and verbal and visual memory.
Precision in results§	Precise
Directness of results‖	Direct
Comparison 2	**Cognitive functioning in people with first-episode bipolar**

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

	disorder vs. people with first-episode schizophrenia.
Summary of evidence	***Memory:*** **Moderate to high quality evidence (medium to large samples, consistent, precise, direct) shows a medium-sized effect of poorer verbal memory in people with first-episode schizophrenia compared to people with first-episode bipolar disorder. Moderate quality evidence (inconsistent) also shows a small effect of poorer working memory. There were no significant differences in digit span or visual memory.** ***Verbal fluency:*** **Moderate to high quality evidence (medium to large samples, consistent, precise, direct) shows a medium-sized effect of poorer verbal fluency in people with first-episode schizophrenia compared to people with first-episode bipolar disorder.** ***Processing speed:*** **Moderate quality evidence (medium to large samples, inconsistent) shows a small effect of poorer processing speed in people with first-episode schizophrenia compared to people with first-episode bipolar disorder.** ***IQ:*** **Moderate to high quality evidence (medium to large samples, consistent, precise, direct) shows a medium-sized effect of poorer premorbid IQ in people with first-episode schizophrenia compared to people with first-episode bipolar disorder. Moderate to low quality evidence (imprecise and inconsistent) also shows a medium-sized effect of poorer current IQ in people with first-episode schizophrenia.** ***Global cognition:*** **Moderate to high quality evidence (large sample, inconsistent, precise, direct) shows a small effect of poorer global cognition in people with first-episode schizophrenia compared to people with first-episode bipolar disorder.** **No differences in attention or reasoning are reported.**
Global cognition	
A significant, small effect of poorer global cognition in people with first-episode schizophrenia compared with first-episode bipolar disorder; 14 studies, N = 1,427, d = 0.28, 95%CI 0.12 to 0.44, $p < 0.001$, I^2 = 48.8%, p = 0.02 Authors report no publication bias.	

No differences were found for males vs. females or younger vs. older patients.
Memory
Significant, small to medium-sized effects of poorer verbal memory, working memory and verbal working memory in people with first-episode schizophrenia compared with first-episode bipolar disorder; All verbal memory tasks: 7 studies, N = 832, *d* = 0.47, 95%CI 0.28 to 0.65, *p* < 0.001, I^2 = 39.5%, *p* = 0.13 Working memory: 8 studies, N = 774, *d* = 0.35, 95%CI 0.11 to 0.59, *p* = 0.005, I^2 = 59.2%, *p* = 0.02 Verbal working memory: 8 studies, N = 774, *d* = 0.33, 95%CI 0.08 to 0.57, *p* = 0.009, I^2 not reported *No significant differences in;* Digit span forwards: 4 studies, N = 435, *d* = 0.18, 95%CI -0.03 to 0.38, *p* = 0.09, I^2 not reported Digit span backwards: 6 studies, N = 536, *d* = 0.13, 95%CI -0.04 to 0.31, *p* = 0.14, I^2 not reported Visual memory: 4 studies, N = 406, *d* = 0.28, 95%CI -0.05 to 0.60, *p* = 0.09, I^2 = 66.2%, *p* = 0.05 Authors report no publication bias. Meta-regression analysis revealed between-group differences in working memory were more significant in studies that included younger people with first-episode schizophrenia. No differences were found for males vs. females.
Processing speed
Significant, small to medium-sized effects of poorer processing speed in people with first-episode schizophrenia compared with first-episode bipolar disorder; All speed tasks: 6 studies, N = 679, *d* = 0.33, 95%CI 0.08 to 0.59, *p* = 0.009, I^2 = 58.9%, *p* = 0.03 TMT A: 3 studies, N = 328, *d* = 0.45, 95%CI 0.23 to 0.68, *p* < 0.001 TMT B: 3 studies, N = 328, *d* = 0.47, 95%CI 0.14 to 0.80, *p* = 0.006 Digit symbol: 3 studies, N = 450, *d* = 0.71, 95%CI 0.36 to 1.06, *p* < 0.001 Authors report no publication bias. No differences were found for males vs. females or younger vs. older patients.
IQ
Significant, medium-sized effects of lower premorbid and current IQ in people with first-episode schizophrenia compared with first-episode bipolar disorder; Premorbid IQ: 7 studies, N = 728, *d* = 0.50, 95%CI 0.30 to 0.69, *p* < 0.001, I^2 = 36.8%, *p* = 0.15 Current IQ: 6 studies, N = 533, *d* = 0.63, 95%CI 0.36 to 0.91, *p* < 0.001, I^2 = 67.9%, *p* = 0.05 Authors report no publication bias. No differences were found for males vs. females or younger vs. older patients.

Fluency	
Significant, medium-sized effects of poorer fluency in people with first-episode schizophrenia compared with first-episode bipolar disorder; All fluency tasks: 7 studies, N = 865, d = 0.50, 95%CI 0.33 to 0.66, $p < 0.001$, I^2 = 22.0%, p = 0.26 Letter: 5 studies, N = 542, d = 0.42, 95%CI 0.24 to 0.60, $p < 0.001$ Category: 3 studies, N = 328, d = 0.77, 95%CI 0.0 to 1.53, p = 0.05 Authors report no publication bias. No differences were found for males vs. females or younger vs. older patients.	
Attention	
No significant differences in attention; 2 studies, N = 101, d = 0.05, 95%CI −0.38 to 0.47, p = 0.83, I^2 = 0%, p = 0.62 Authors report no publication bias. No differences were found for males vs. females or younger vs. older patients.	
Reasoning	
No significant differences in reasoning; 2 studies, N = 218, d = 0.23, 95%CI −0.09 to 0.56, p = 0.16, I^2 = 26.3%, p = 0.24 Authors report no publication bias. No differences were found for males vs. females or younger vs. older patients.	
Consistency in results	Consistent for verbal memory, premorbid IQ, fluency, attention and reasoning. Inconsistent for global cognition, working memory, visual memory, psychomotor speed, and current IQ.
Precision in results	Precise for global cognition, verbal memory, working memory, visual memory, digit span, psychomotor tasks, TMT A, premorbid IQ, and fluency (all and letter). Imprecise for TMT B, digit symbol, current IQ, category fluency, attention, and reasoning.
Directness of results	Direct

Trotta A, Murray RM, MacCabe JH

Do premorbid and post-onset cognitive functioning differ between schizophrenia and bipolar disorder? A systematic review and meta-analysis

Psychological Medicine 2015; 45: 381-94

View review abstract online

Comparison	**General cognitive functioning in people with first-episode bipolar disorder compared to controls vs. general cognitive functioning in people with chronic bipolar disorder compared to controls.**
Summary of evidence	**Moderate quality evidence (medium to large samples, some inconsistency, precise, direct) suggests a small effect of poorer general cognition in people with first-episode bipolar disorder compared to controls. This effect was smaller than the medium-sized effect found in chronic patients.**
General cognition	
Small, significant effect of more general cognitive impairment in people with first-episode bipolar disorder than in controls; First episode: 3 studies, N = 390, SMD = -0.277, 95%CI -0.510 to -0.044, p = 0.020, I^2 = 15.1%, p = 0.308 *The effect was medium-sized in chronic patients;* Chronic patients: 16 studies, N = 1,789, SMD = -0.691, 95%CI -0.793 to -0.588, p < 0.0001, I^2 = 82.4%, p < 0.0001	
Consistency in results	Consistent for first-episode, inconsistent for chronic patients.
Precision in results	Precise
Directness of results	Direct

Explanation of acronyms

CI = confidence interval, *d* = Cohen's *d* standardised mean difference, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), N = number of participants, *p* = probability of rejecting a null hypothesis of no differences between groups, SMD = standardised mean difference

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. Standardised 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. Less than 0.4 represents a small effect, around 0.5 a medium effect, and over 0.8 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

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

between variables. They can provide an indirect indication of prediction, but do not 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. Trotta A, Murray RM, MacCabe JH (2015): Do premorbid and post-onset cognitive functioning differ between schizophrenia and bipolar disorder? A systematic review and meta-analysis. *Psychological Medicine* 45: 381-94.
4. Bora E, Pantelis C (2015): Meta-analysis of Cognitive Impairment in First-Episode Bipolar Disorder: Comparison With First-Episode Schizophrenia and Healthy Controls. *Schizophrenia Bulletin* 41: 1095-104.
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*.