

Treatment resistance

Introduction

Medications provide symptom respite and improvement in quality of life for many people with bipolar disorder. However, for some, medications do not provide adequate relief from symptoms. A broad definition of treatment resistance includes those patients whose symptoms have not responded adequately to medications, or only partially responded after four or more weeks of treatment with appropriate doses.

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 bipolar or 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/or comprehensive review was included. Reviews with pooled data have been given priority 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-analysis¹. 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 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 one systematic review that met inclusion criteria³.

- Moderate to low quality evidence suggests non-response to antipsychotic medications in people with acute mania is around 41%, and non-remission is around 65%.
- Early non-response in weeks 1 and 2 (defined by cutoff scores ranging from a $\leq 10\%$ to a $\leq 50\%$ reduction in symptoms compared to baseline on the Youth Mania Rating Scale) significantly predicted non-response ($\leq 0\%$ symptom reduction) and non-remission (YMRS score higher than 8) in week 3, regardless of antipsychotic characteristics.

Treatment resistance

Welten CCM, Koeter MWJ, Wohlfarth TD, Storum JG, Van Den Brink W, Gispen-De Wied CC, Leufkens HGM, Denys DAJP

Early non-response in the antipsychotic treatment of acute mania: A criterion for reconsidering treatment? Results from an individual patient data Meta-Analysis

Journal of Clinical Psychiatry 2016; 77: e1117-e23

[View review abstract online](#)

Comparison	Relationship between early non-response to antipsychotic treatment and treatment failure in people with bipolar disorder with acute mania.
Summary of evidence	Moderate to low quality evidence (large sample, unable to assess consistency or precision, direct) suggests early non-response in weeks 1 and 2 (defined by cutoff scores ranging from a $\leq 10\%$ to a $\leq 50\%$ reduction in symptoms compared to baseline on the Youth Mania Rating Scale) significantly predicted non-response ($\leq 0\%$ symptom reduction) and non-remission (YMRS score higher than 8) in week 3, regardless of antipsychotic characteristics. Non-response rates at study end were around 41% and non-remission around 65%.
Treatment failure	
<p>10 trials, N = 1,243</p> <p>Rates of non-response at study end point: 40.9%, 95%CI, 38.2% to 43.6%</p> <p>Rates of non-remission at study end point: 65.3%, 95%CI, 62.0% to 68.6%</p> <p>Early non-response in weeks 1 and 2 (defined by cutoff scores ranging from a $\leq 10\%$ to a $\leq 50\%$ reduction in symptoms compared to baseline on the Youth Mania Rating Scale) significantly predicted non-response ($\leq 0\%$ symptom reduction) and non-remission (YMRS score higher than 8) in week 3, regardless of antipsychotic characteristics.</p>	
Consistency in results[†]	Unable to assess, no measure of consistency is reported.
Precision in results[§]	Unable to assess, no measure of precision is reported.
Directness of results	Direct

Treatment resistance

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 small⁴.

† 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 effect⁴.

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 ⁵. InOR stands for logarithmic OR where a InOR 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

Treatment resistance

between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforeseen 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.

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 relaxed⁶.

‡ 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 formula⁴;

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

|| 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.

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous

Treatment resistance

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADE Working Group (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Welten CCM, Koeter MWJ, Wohlfarth TD, Storsum JG, Van Den Brink W, Gispen-De Wied CC, et al. (2016): Early nonresponse in the antipsychotic treatment of acute mania: A criterion for reconsidering treatment? Results from an individual patient data Meta-Analysis. *Journal of Clinical Psychiatry* 77: e11117-e23.
4. Cochrane Collaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
5. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
6. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. Version 3.2 for Windows.