TECHNICAL COMMENTARY

NeuRA Discover. Conquer. Cure.
BIPOLAR DISORDERS LIBRARY

Sleep deprivation

Introduction

Sleep deprivation, or 'wake therapy', involves being deprived of total sleep for approximately 36 hours straight, or partially sleeping for only 3 to 4 hours followed by 20 to 21 hours of wakefulness. This therapy has been associated with rapid improvements in depressed mood but may also trigger mania. Studies have used various treatment formats in the number, timing, and duration of sleep deprivation cycles to determine which works best for people with bipolar depression.

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 a bipolar disorder. Reviews were identified by searching the databases MEDLINE, EMBASE, and PsycINFO. When multiple copies of review topics were found, the most recent and/or 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 (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, 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 quality evidence shows total sleep deprivation can improve depression in around half of people receiving the treatment. It is particularly effective when accompanied by mood stabilizers or antidepressants. There were no further improvements when adding sleep phase advance or light treatment, or when adding more than one sleep deprivation exposure.
- Moderate to low quality evidence finds a medium-sized improvement in depression with total sleep deprivation plus medication compared to medication alone. This effect may remain for up to 4 weeks post-treatment. Adding mood medication after 10 days of total sleep deprivation treatment improved depression and increased rates of remission by 3 months. Around 4% of patients had an episode of mania or hypomania with total sleep deprivation.

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

TECHNICAL COMMENTARY

Sleep deprivation

Gottlieb JF, Goel N, Chen S, Young MA

Meta-analysis of sleep deprivation in the acute treatment of bipolar depression

Acta Psychiatrica Scandinavica 2021; 143: 319-27

View review abstract online

Comparison	**Pre-post total sleep deprivation in people with bipolar depression.** **Note: one included study used partial sleep deprivation (3hrs in bed between 11pm and 2am).**
Summary of evidence	**Moderate quality evidence (medium-sized sample, inconsistent, precise, direct) suggests total sleep deprivation may improve depression in around half of people with bipolar disorder, particularly if accompanied by mood stabilizers or antidepressants. There were no differences in response rates with adjunctive or no adjunctive chronotherapeutic treatment (sleep phase advance or light treatment), one or three sleep deprivation exposures or bipolar disorder type.**

Depression

Around 50% of people with bipolar depression improved after treatment with total sleep deprivation;

15 studies, N = 384, mean response rate = 47.6%, 95%CI 36.0% to 59.5%, I^2 = 70%

People receiving sleep deprivation plus mood medications (mostly lithium or antidepressants) showed greater response than people receiving sleep deprivation only (59.4% vs. 27.4%, $p < 0.001$).

There were no significant differences in response rates according to having adjunctive chronotherapeutic treatment (sleep phase advance or light treatment = 58.3% vs. none = 38.7%, $p = 0.114$), the number of exposures (one = 56.8% vs. three = 45.1%, $p = 0.278$), type of response criterion (percentage = 54.3% vs. score = 43.6%, $p = 0.415$), or bipolar type (I = 48.6% vs. I and II = 46.6%).

Risks	Not reported
Consistency‡	Inconsistent
Precision§	Appears precise
Directness‖	Direct

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

Ramirez-Mahaluf JP, Rozas-Serri E, Ivanovic-Zuvic F, Risco L, Vöhringer PA

Effectiveness of sleep deprivation in treating acute bipolar depression as augmentation strategy: A systematic review and meta-analysis

Frontiers in Psychiatry 2020; 11: 9

View review abstract online

Comparison	**Total sleep deprivation plus mood medication in people with bipolar depression vs. medication alone.**
Summary of evidence	**Moderate to low quality evidence (very small samples, mostly consistent, imprecise, direct) finds a significant, medium-sized effect of improved depression after total sleep deprivation plus medication versus medication alone, which remained, but was not significant at 4 weeks follow-up.** **Adding mood medication after 10 days of total sleep deprivation treatment was associated with a significant improvement in depressive symptoms compared to total sleep deprivation alone Adding mood medication to total sleep deprivation treatment was also associated with increased rates of remission by 3 months. Around 4% of patients had an episode of treatment-emergent mania or hypomania with total sleep deprivation.**
Depression	
A significant, medium-sized effect of improved depression after total sleep deprivation plus medication vs. medication alone, which remained, but was not significant at follow-up; Post-treatment: 2 studies, N = 59, *d* = -0.58, 95%CI -1.13 to -0.04, *p* = 0.03, I^2 = 0% 2 weeks: 2 studies, N = 59, *d* = -0.41, 95%CI -0.95 to 0.12, *p* = 0.13, I^2 = 0% 4 weeks: 2 studies, N = 59, *d* = -0.45, 95%CI -0.99 to 0.09, *p* = 0.10, I^2 = 0% *Adding mood medication after 10 days of total sleep deprivation treatment was associated with a significant improvement in depressive symptoms compared to total sleep deprivation alone;* 4 studies, N = 137, *d* = -0.89, 95%CI -1.39 to -0.40, *p* < 0.001, I^2 = 48% *Adding mood medication to total sleep deprivation treatment was also associated with increased rates of remission by 3 months;* 3 studies, N = 110, OR = 2.36, 95%CI 0.93 to 3.78, *p* < 0.001, I^2 = 46%	
Risks	4.35% of patients given medication plus total sleep deprivation, 4.17% of patients given total sleep deprivation alone, and 0% of patients given mood medication alone, had an episode of treatment-emergent mania or hypomania.
Consistency	Consistent for all analyses apart from adding mood medication.

Sleep deprivation

Precision	Imprecise
Directness	Direct

Explanation of acronyms

CI = confidence interval, *d* = Cohen's d standardized 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, OR = odds ratio, *p* = statistical probability of obtaining that result ($p < 0.05$ generally regarded as significant), vs. = versus

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

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 confirm causality due to possible and often

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

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. Gottlieb JF, Goel N, Chen S, Young MA (2021): Meta-analysis of sleep deprivation in the acute treatment of bipolar depression. *Acta Psychiatrica Scandinavica* 143: 319-27.
4. Ramirez-Mahaluf JP, Rozas-Serri E, Ivanovic-Zuvic F, Risco L, Vöhringer PA (2020): Effectiveness of sleep deprivation in treating acute bipolar depression as augmentation strategy: A systematic review and meta-analysis. *Frontiers in Psychiatry* 11: 9.
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