

BMJ Open Does improving sleep lead to better mental health? A protocol for a meta-analytic review of randomised controlled trials

Alexander J Scott,¹ Thomas L Webb,² Georgina Rowse³

To cite: Scott AJ, Webb TL, Rowse G. Does improving sleep lead to better mental health? A protocol for a meta-analytic review of randomised controlled trials. *BMJ Open* 2017;**7**:e016873. doi:10.1136/bmjopen-2017-016873

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2017-016873>).

Received 16 March 2017

Revised 18 July 2017

Accepted 25 July 2017



CrossMark

¹School of Health and Related Research, Sheffield University, Sheffield, UK

²Department of Psychology, Sheffield University, Sheffield, UK

³Clinical Psychology Unit, Sheffield University, Sheffield, UK

Correspondence to

Dr Alexander J Scott;
alex.scott@sheffield.ac.uk

ABSTRACT

Introduction Sleep and mental health go hand-in-hand, with many, if not all, mental health problems being associated with problems sleeping. Although sleep has been traditionally conceptualised as a secondary consequence of mental health problems, contemporary views prescribe a more influential, causal role of sleep in the formation and maintenance of mental health problems. One way to evaluate this assertion is to examine the extent to which interventions that improve sleep also improve mental health.

Method and analysis Randomised controlled trials (RCTs) describing the effects of interventions designed to improve sleep on mental health will be identified via a systematic search of four bibliographic databases (in addition to a search for unpublished literature). Hedges' *g* and associated 95% CIs will be computed from means and SDs where possible. Following this, meta-analysis will be used to synthesise the effect sizes from the primary studies and investigate the impact of variables that could potentially moderate the effects. The Jadad scale for reporting RCTs will be used to assess study quality and publication bias will be assessed via visual inspection of a funnel plot and Egger's test alongside Orwin's fail-safe *n*. Finally, mediation analysis will be used to investigate the extent to which changes in outcomes relating to mental health can be attributed to changes in sleep quality.

Ethics and dissemination This study requires no ethical approval. The findings will be submitted for publication in a peer-reviewed journal and promoted to relevant stakeholders.

PROSPERO registration number CRD42017055450.

INTRODUCTION

Difficulties sleeping and mental health problems are both public health concerns in their own right, with each having a substantive impact on both individuals and society as a whole.^{1–4} However, sleep and mental health go hand-in-hand, with many, if not all, mental health problems being associated with problems sleeping.^{5–7} Traditionally, sleep problems have been viewed as a consequence of mental health problems. Although this is not contested, evidence also suggests that

Strengths and limitations of this study

- The proposed review should provide reliable evidence on the effect of interventions designed to improve sleep on outcomes reflecting mental health.
- The findings of the proposed review will further elucidate the nature of the relationship between sleep and mental health.
- The GRADE (Grading of Recommendations, Assessment, Development and Evaluation) system will be used to assess the strength of the evidence base and allow members of the public, researchers and clinicians to judge the quality of the available evidence.
- The proposed review will include a diverse range of interventions and target problems and so might lead to a heterogeneous group of studies. However, to mitigate this, moderation analysis will be used to investigate specific factors that might influence the effect of sleep improvement on mental health.

problems sleeping can contribute to the formation of new mental health problems^{8–10} and to the maintenance of existing ones.^{11–13}

In other words, sleep is now thought to have a bidirectional relationship with mental health, with problems sleeping likely to influence both the onset and trajectory of a variety of mental health difficulties. Having said this, although a number of empirical studies have manipulated sleep and examined the impact of doing so on outcomes related to mental health, to date there has not been a systematic review of these studies. Consequently, the magnitude of the effect of (changes in) sleep on mental health problems is difficult to estimate and has not been compared between different mental health outcomes and other factors that might influence the effect (eg, across different groups of participants, research designs and approaches to intervention).

The potential for a causal relationship between sleep and mental health also raises

an exciting prospect; namely, that interventions designed to improve sleep could also improve mental health. Providing a definitive answer to this question would have important implications for clinicians, researchers and members of the public alike. From a practical perspective, if interventions designed to improve sleep can change mental health outcomes, then they may be a useful tool for tackling mental health difficulties. Indeed, interventions designed to improve sleep can often be delivered remotely, in self-help and group formats, and/or at little cost through the internet.^{14–18} For example, a meta-analysis by Ho *et al* reported that self-help interventions based on the principles of cognitive behavioural theory (CBT) for insomnia (termed CBTi) had medium to large effects on the symptoms of insomnia.¹⁸

Current evidence on the relationship between sleep and mental health

The relationship between sleep and mental health is well documented, with numerous reviews testifying to a robust link between the two.^{6–8 19–24} However, the majority of these reviews have focused on primary studies with correlational research designs. That is, they (1) measure associations between variables at a single time point (ie, cross-sectional designs); (2) measure associations between variables at multiple time points (ie, longitudinal designs) or (3) compare the typical sleep profiles of those with mental health difficulties to those without.^{6 7 25 26} Unfortunately, cross-sectional designs simply tell us that variables are associated in some way. It is impossible to determine whether sleep causes mental health problems, mental health problems cause difficulties in sleeping or whether the effect is bidirectional in nature.

Longitudinal studies, although still correlational in nature, are better able to elucidate causality than their cross-sectional counterparts. However, only a handful of reviews have provided evidence on the relationship between sleep (at one point in time) and mental health outcomes (measured later). Furthermore, all of these have focused on depression.^{8 24 27 28} For example, Baglioni *et al*⁸ meta-analysed 21 studies that investigated the longitudinal associations between insomnia and depression. Baglioni *et al* reported that people with insomnia had a twofold risk of developing depression compared with people who did not experience difficulties in sleeping. Longitudinal designs are also still susceptible to the ‘third variable problem’.^{29–31} Namely, that a third, unmeasured variable (eg, having young children) could cause both sleep difficulties and mental health problems. In summary, correlational designs are not a valid way of disentangling the relationship between problems sleeping and mental health.

Some reviews have assessed the impact of interventions designed to improve sleep on mental health outcomes.^{17 18 23 24 32–36} However, for a number of reasons, even these reviews do not permit us to draw robust conclusions as to the causal impact of sleep quality on mental health outcomes. First, these reviews often include

interventions that have not successfully manipulated sleep (ie, studies in which there was no significant impact of the intervention on sleep outcomes). Such studies do not tell us anything about the relationship between sleep and mental health other than that it can be difficult to improve sleep. Second, the focus of extant reviews has been on improving sleep, with the measurement of mental health outcomes being secondary and typically limited to depression and anxiety. Consequently, the effect of improving sleep on other mental health problems is currently unclear.

Finally, to our knowledge, to date there has been no attempt to investigate variables that influence—or moderate—the impact of interventions that improve sleep on mental health. However, interventions designed to improve sleep are likely to vary in their content and delivery, and such variables may influence how effective they are (or appear to be) in improving sleep and/or mental health outcomes. Furthermore, variables related to the nature of the sample (eg, age, severity of symptoms, nature of the mental health problem) and methodological features of the study (eg, self-report vs objective assessment of the outcome variables) are likely to influence the apparent effect of the intervention. It is therefore crucial that the impact of such variables is systematically examined across the extant evidence base in order to draw reliable and valid conclusions about the impact of changes in sleep on outcomes pertaining to mental health.

The proposed review

A number of primary research studies have experimentally manipulated sleep (typically via some sort of psychological intervention) and then measured mental health outcomes. However, as described above, these individual studies have, to our knowledge, never been integrated in a manner that allows the magnitude of the effect of sleep quality on mental health outcomes to be estimated. Therefore, it is currently difficult to: (1) draw firm conclusions about the relationship between sleep and various mental health problems and (2) recommend with any confidence that mental health problems might be tackled using interventions that have been designed to improve sleep. Furthermore, to date there has been no attempt to understand the factors that influence, or moderate, the effect of improvements in sleep on mental health. As a consequence, it is currently unclear whether and how the content and nature of the intervention(s), target sample and mental health problem, and methodological features of the primary study influence the effects of interventions designed to improve sleep on mental health outcomes.

Objectives

The proposed review therefore has two broad objectives: (1) to synthesise and quantify the effect of interventions that improve sleep on outcomes reflecting mental health and (2) to explore variables that moderate the effect

of interventions targeting sleep on outcomes reflecting mental health.

METHOD AND ANALYSIS

This protocol has been prepared in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P, see online supplementary material 1) checklist.³⁷

Outcomes and prioritisation

Measuring improvements in sleep

The concept of 'improved sleep' is multifaceted and can mean different things to different people.^{38–40} Indeed, many specific sleep problems are tied to mental health in unique ways and often have their own unique measures. For example, the experience of nightmares has been found to be associated with post-traumatic stress disorder (PTSD),⁴¹ as measured using specific outcome measures such as dream diaries^{42 43} or the Clinician-Administered PTSD Scale.⁴⁴ Consequently, one challenge for the proposed review is to ensure that all of the primary studies assess a similar notion of sleep improvement. To achieve this, the proposed review will require that the primary studies report a measure that reflects the overall quality of sleep experienced by participants. Broadly speaking, sleep quality consists of: (1) sleep continuity (eg, sleep onset, sleep maintenance and number of awakenings) and (2) daytime impact (eg, the extent to which the person feels refreshed on waking and throughout the day).^{39 40}

Sleep quality can be measured using both self-report and objective indices. For example, the Pittsburgh Sleep Quality Index⁴⁵ is widely recognised as the 'gold standard' self-report measure of sleep quality and consists of 19 items measuring seven aspects of sleep quality (namely, subjective sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbances, use of sleep medication and daytime dysfunction). The 'gold standard' for objectively measuring sleep is accepted to be polysomnography⁴⁶; a technique that monitors multiple biophysiological parameters and directly records aspects of sleep quality including sleep onset and sleep maintenance (for a review, see Vaughn and Giallanza⁴⁷). As such, the proposed review will include both self-report and objective indices of sleep quality, but will also seek to compare effect sizes between different measures in an effort to empirically examine the extent to which the nature of the measures influences the apparent effect of the interventions.

Measuring mental health

Measuring mental health is also complex and multifaceted, with a range of different outcomes which differ both in their administration and interpretation. Some studies will report a general measure assessing a specific diagnostic category (eg, a measure of the severity of depression). For example, the Beck Depression Inventory II is

a 21-item self-report measure designed to assess multiple facets of depression including mood, pessimism, self-dislike, loss of appetite and social withdrawal, with higher scores indicating more severe depression.⁴⁸ Other studies might assess a single symptom or problem. For example, the Green Paranoid Thoughts Scale measures paranoid thoughts,⁴⁹ an experience that is associated with, but is not limited to, psychosis spectrum disorders.^{50 51} Finally, some studies may report the effects of interventions designed to improve sleep on global measures of mental health. For example, the Clinical Global Impressions Severity scale⁵² asks clinicians to use their clinical experience to rate how mentally ill their client has been over the last week, on a scale ranging from 1—normal to 7—among the most extremely ill patients.

Measures assessing aspects of mental health can either be: (1) self-reported by the participant or (2) completed on behalf of the participant by a clinician or other independent rater. Both self-report and independently rated outcome measures will be included in the proposed review; however, as above, we will compare effect sizes between different measures in an effort to empirically examine the extent to which the nature of the measure(s) influences the apparent effect of the interventions.

Eligibility criteria

Inclusion criteria

In order to be included in the proposed review, the primary studies need to:

1. Randomly allocate participants to either an experimental group that receives an intervention that is designed to improve sleep or a comparison group.
2. Report a statistically significant improvement on a measure of sleep quality in at least one follow-up point among participants in the experimental group as compared with those in the comparison group.
3. Include a measure of mental health subsequent to the measure of sleep quality.
4. Report sufficient data for us to be able to compute effect sizes reflecting the impact of the intervention on (1) sleep quality and (2) mental health. Where sufficient data is not reported, we will contact the authors and request further data. However, if this is not provided, then the study will not be included in the review.
5. Be written in English or be able to be translated using available translation resources.

Exclusion criteria

The aim of the proposed review is to be as inclusive as possible and address potential differences between the primary studies (eg, differences in the nature of the intervention or the mental health problem under consideration) using moderation analysis. Therefore, we will not restrict the type of intervention (eg, psychological and pharmacological), publication status, nature of the comparison condition or sample (ie, interventions directed towards adults, children and adolescents will all

be eligible). However, in order to ensure that we can reliably and validly assess the independent contribution of changes in sleep on mental health outcomes among adult populations, studies with the following characteristics will be excluded:

1. Studies where the intervention contains elements that specifically target a mental health problem alongside improving sleep (eg, an intervention that provides CBT for anxiety alongside efforts to improve sleep).
2. Studies that recruit children and young people (ie, under the age of 18 years old).
3. Studies adopting a pre-post design (ie, within participant designs).

Information sources

The proposed review will use a combination of search techniques and sources in order to identify potential studies. First, we will search MEDLINE (1946 to present), Embase (1974 to present), PsycINFO (1967 to present) and The Cochrane Library (1898 to present) using the Cochrane Highly Sensitive Search Strategy⁵³ to identify RCTs that include terms relating to sleep quality and/or sleep disorders and mental health (see table 1 for a list of the proposed search terms). The search strategy has been developed in collaboration with a health sciences librarian specialising in systematic search procedures and will be used to search each database (see online supplementary material 2 for an example search strategy). Second, the reference lists of extant reviews of the relationship between sleep and mental health (eg, those cited in the introduction) will be searched for any potential articles. Third, a search for any unpublished or ongoing studies will be conducted by searching online databases including White Rose Online, The National Research Register, WHO approved clinical trial databases (eg, ISRCTN) and PROSPERO. Finally, the authors of articles deemed eligible for inclusion will be contacted and asked if they are aware of any unpublished research that may be eligible for inclusion in the review.

Data management

All records will be stored in the reference management software Endnote, and we will follow PRISMA guidelines for the selection of studies for meta-analysis.⁵⁴ Specifically, when the pool of potential studies has been identified, we will remove duplicates and initially screen each record based on the title and abstract and exclude clearly ineligible studies. Following this initial screening, the full-text versions of each article will be reviewed in detail and cross-referenced against the inclusion and exclusion criteria. The flow of articles through the review, including the reasons for excluding studies, will be documented in a PRISMA flow chart.

Data extraction

Data will be recorded on a standardised data extraction form and a manual will accompany this form and detail each variable to be extracted alongside definitions and

Table 1 Search terms that will be used to identify RCTs of interventions designed to improve sleep on outcomes pertaining to mental health

HSSS for RCTs†	Sleep	Mental health
Randomized controlled trial	Sleep*	“Psychological health”
Controlled clinical trial	“Circadian rhythm”	“Mental”
Randomized	Insomnia	Psychiat*
Placebo	Hypersomnia	Affect*
Drug therapy	Parasomnia	Depress*
Randomly	Narcolepsy	Mood
Trial	Apnea	Stress
Groups	Apnoea	Anxi*
	Nightmare*	Phobi*
	“Restless legs syndrome”	“Obsessive compulsive disorder”
		OCD
		PTSD
		“Post-traumatic stress disorder”
		Psychos*
		Psychotic
		Schiz*
		Bipolar
		Hallucination*
		Delusion*
		“Eating disturbance”
		Anorexia
		Bulimia
		“Binge eating”

Studies will need to include at least one search term from each of the filter above in the title, abstract or keywords, for consideration for inclusion in the review.* indicates that variants of the word after the asterisk will be searched for (eg, depress* will search for depressive, etc)

†The highly sensitive search strategy (HSSS) is more than just a keyword search, rather it encompasses search techniques and strategies.⁵³

OCD, obsessive compulsive disorder; PTSD, post-traumatic stress disorder; RCT, randomised controlled trial.

examples (see online supplementary materials 3 and 4). Two reviewers will pilot the data extraction forms and manual on three articles in order to ensure that there are no systematic problems or difficulties in coding any of the variables. After this, the data will be extracted from the full set of studies by one reviewer. A second member of the review team will second code a subset of the included articles (at least 10%) and levels of agreement will be calculated (the subset of articles for second coding will be randomly selected using a computer generated algorithm). Any disagreements will be resolved through discussion, with a third member of the review team acting

Table 2 Variables to be extracted for moderation analysis (where available, see supplementary material 3 for detailed variable categories and levels)

Source characteristics	Sample characteristics	Design characteristics	Intervention characteristics
Publication status	Age	Method of recruitment	Size of the effect on sleep
Publication year	Gender	Nature of comparison group(s)	Duration
Journal impact factor	Type of mental health problem(s)	Attrition/dropout rate	Theoretical basis
	Type of sleep problem(s)	Timing of follow-up	Mode of delivery
	Clinical status of mental health	Nature of outcome measure(s)	Adherence
	Clinical status of sleep problem	Type of analysis	
	Comorbidity	Adjusted versus unadjusted data	
	Concurrent medication use	Study quality	
	Concurrent psychological help		

as an arbiter for any outstanding disagreements. The review team will extract meta-data pertaining to source characteristics (eg, publication status and year), as well as data relating to the characteristics of the sample (eg, age, type of mental health problem), the study (eg, the nature of the comparison group, length of follow-up) and characteristics of the intervention (eg, theoretical basis, mode of delivery). [Table 2](#) provides an overview of the potential moderators that we propose to code and examine, and online supplementary material 3 provides detail on specific moderator levels and categories.

Proposed analysis

Review Manager V.5.3 (Cochrane Collaboration)⁵⁵ will be used to compute Hedges' g (and associated 95% CIs) using the means and standard deviations for each measure of sleep quality and mental health reported in studies comparing these outcomes between an intervention group (ie, a group receiving an intervention that improves sleep) and a comparison group (eg, wait-list, placebo, treatment as usual)¹. Where means and standard deviations are not available, we will compute effect sizes by converting relevant summary statistics (eg, F values from an analysis of variance, testing the impact of an intervention on relevant outcomes) using Lyons Morris' meta-analysis calculator.⁵⁶ The effect of the interventions on sleep quality will be assessed using data from the first available follow-up point that reports a statistically significant difference in sleep quality between the intervention and comparison conditions. The effect of the interventions on outcomes pertaining to mental health will be

assessed at the longest follow-up point available, whether the effect at this point is statistically significant or not (and we will investigate the effect of follow-up duration on outcomes using moderation analysis). This strategy will provide a stringent test of the effect of the interventions on outcomes pertaining to mental health (in the sense that any changes need to have been maintained over time) and also enable us to investigate whether the impact of the interventions on outcomes is mediated by changes in sleep quality that precede the impact on outcomes pertaining to mental health (this proposed analysis is discussed in the 'mediation analysis' section).

Where studies report multiple outcome measures under one diagnostic category (eg, several measures of depression or sleep quality), the effect sizes will be computed for each outcome and meta-analysed in their own right to form one overall effect for inclusion in the main analysis. For example, we would compute two effect sizes reflecting sleep quality if a study reported the effects of an intervention on the Pittsburgh Sleep Quality Index⁴⁵ and the Insomnia Severity Index⁵⁷ (ie, one effect size for each measure of sleep quality) and then average them before inclusion in the main dataset. This procedure capitalises on the information that is available, while retaining the independence of effect sizes which is central to the validity of meta-analysis.⁵⁸

The sample-weighted average effect size (g_i) will be computed using a random effects model as studies are likely to be 'different from one another in ways too complex to capture by a few simple study characteristics'.⁵⁹ Following Cohen's⁶⁰ recommendations, $g=0.20$ will be taken to represent a 'small' effect size, $g=0.50$ a 'medium' effect size and $g=0.80$ a 'large' effect size. We will use these qualitative indices to interpret the findings. Publication bias will be assessed via visual inspection of a funnel plot and Egger's test.⁶¹ Finally, Orwin's⁶² formula will be used to determine the fail-safe n (ie, the number of studies producing a null effect that would be needed to reduce the overall effect of interventions that improve sleep on outcomes relating to mental health to a trivial effect size).

¹Wherever available, data that have been adjusted for baseline differences between groups will be used to compute effect sizes. However, if this information is not reported then we will use the unadjusted data to compute the effect sizes. We will also seek to compute effect sizes using the data from intention-to-treat (ITT) analyses where they are reported. Subscripts will be added to the table reporting the effect sizes derived from the primary research studies in order to identify how each effect size was computed and also to compare outcomes between studies that report adjusted versus unadjusted statistics and ITT analyses versus non-ITT analyses.

Heterogeneity, bias and study quality

The I^2 statistic will be used to assess the heterogeneity of effect sizes across the primary studies.⁶³ The quality of each individual study included in the proposed review will be assessed using the Jadad scale for reporting RCTs.⁶⁴ The Jadad scale assesses three key areas of methodological quality that potentially lead to bias—namely, randomisation, blinding and the flow of participants through the study. In order to assess these areas, raters will be asked to answer three questions: (1) ‘Was the study described as randomized (ie, does it include words such as randomly, random, and randomization)?’; (2) ‘Was the study described as double blind?’; and (3) ‘Was there a description of withdrawals and dropouts?’. Scores on the Jadad scale range from 0 to 5, with higher scores indicating a lower risk of bias (and therefore higher methodological quality). The Jadad scale for reporting RCTs has been extensively used as a measure of the methodological quality of RCTs (having received over 7500 citations to date) and has been recommended as the most reliable and valid scale for assessing the quality of RCTs in a review of 21 measures.⁶⁵ Finally, the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system^{66 67} will be used to assess the quality of the body of evidence as a whole and the extent to which it can and should be used to inform clinical recommendations.

Moderation analysis

Moderation analyses will be used to identify variables that influence the effect of interventions that improve sleep on both mental health. Many of these variables and their subcategories are outlined in [table 2](#) (for more detail see online supplementary material 3); however, we are keen to be flexible and responsive to the literature as the search develops. Imposing an exhaustive coding structure a priori without knowledge of the primary studies included in the review may result in an unsuitable structure that does not accurately reflect the nature of the included studies. Consequently, [table 2](#)/online supplementary material 3 is not intended to provide an exhaustive list of moderators and we are open to considering additional moderators and categories as the search and data extraction develops. However, in order to ensure that the reader is clear on what analyses were preplanned, we will label any analyses that are not prespecified in this protocol as exploratory in the final report. Moderation analysis will be undertaken to explore the effect of variables relating to the nature of the focal sample, the methodological design and intervention characteristics across all studies within the main meta-analyses.

We will require a minimum of $k=3$ studies representing each moderator level category in order to conduct moderation analysis (eg, to investigate the effect of outcome type on effect sizes, we will require data from at least three studies using self-report outcomes and at least three studies using clinician completed outcomes). For continuous moderators (eg, age, publication year, study quality), sample weighted meta-regression will be

used to investigate the impact of the moderator on effect sizes. For example, the quality of a given study, assessed using the Jadad scale,⁶⁴ will be used as the independent variable in a sample-weighted meta-regression, with the effect sizes representing the effect of the interventions on outcomes pertaining to mental health used as the dependent variable. For categorical variables (eg, self-report vs clinician rated outcomes, the nature of the comparison condition), the sample-weighted average effect size (g_i) and associated standard errors will be computed for each level of the moderator and then the Q statistic will be used to assess if the effect sizes are significantly different. For example, effect sizes based on clinician rated measures of mental health (eg, the Clinical Global Impressions Severity Scale⁵²) will be compared with effect sizes based on self-report measures of mental health (eg, the Depression, Anxiety and Stress Scale⁶⁸).

Mediation analysis

Mediation analysis will be used to investigate the extent to which changes in mental health can be attributed to changes in sleep. These analyses will include all studies that report the correlation between (changes in) sleep quality and (changes in) mental health outcomes (the correlation between the intervention and sleep quality and mental health outcomes, respectively, will be computed by converting the sample-weighted average effect of the interventions on these outcomes into effect size r). These (sample-weighted, average) correlations will be entered using the matrix function into SPSS (version 23) to permit analysis as if they resulted from a primary dataset. In line with Kenny, Kashy and Bolger’s⁶⁹ recommendations, we will then conduct four multiple regressions in order to investigate mediation. These regressions will test: (1) the effect of the independent variable (ie, the intervention) on the dependent variable (ie, outcomes reflecting mental health); (2) the effect of the independent variable on the putative mediator (ie, outcomes reflecting sleep quality); (3) the effect of the mediating variable on the dependent variable and (4) the simultaneous effect of the independent variable and the mediator on the dependent variable, respectively. If the effect of the interventions on mental health can be attributed to changes in the quality of sleep, then the impact of the interventions on outcomes pertaining to mental health should be significantly reduced when the effect of the interventions on sleep quality is statistically controlled.

Ethics and dissemination

As the proposed research is a meta-analytic review of primary studies, no ethical approval is required. We have registered the proposed review on the PROSPERO database (CRD42017055450) in order to adhere to the principles of open research. Following completion of the review, we will submit the findings for publication in a peer-reviewed academic journal and attend conferences and dissemination events with stakeholders wherever possible.

DISCUSSION

The proposed review will use meta-analysis alongside moderator and (meta)mediation analyses to (1) quantify the effect of interventions that improve sleep on mental health outcomes, (2) test whether any effect of the interventions on these outcomes is mediated by changes in sleep quality and (3) explore variables that potentially moderate the effect of the interventions targeting sleep on mental health outcomes. The proposed review has a number of strengths that we believe will make a substantive contribution. First, the review will be inclusive and investigate the effect of improving sleep on a wide range of mental health problems. Second, the review will further elucidate our understanding of the causal relationship between sleep and mental health by including only studies that successfully manipulate sleep and by conducting a mediation analysis to investigate whether any changes in mental health can be attributed to changes in sleep. Finally, the GRADE system will be used to assess the strength of the evidence base^{66 67} which should allow members of the public, researchers and clinicians to quickly access the available evidence and judge its quality.

Despite the strengths of the proposed review, however, the wide range of interventions and target problems that are likely to be addressed by the primary research studies may lead to a relatively heterogeneous group of studies (and thus, potentially effect sizes) which may lead to concerns that we are not comparing 'like with like' (cf. the problem of mixing apples and oranges⁷⁰) and limit the extent to which the findings can be generalised to a specific population (eg, to those with depression). However, to mitigate these concerns we will use moderation analysis to investigate specific factors that might influence the effect of improvements in sleep on mental health and to estimate the sample-weighted average effect sizes for different types of interventions and on different mental health outcomes. Our hope is that these analyses prove informative, both in understanding mental health problems (ie, for which mental health problems can changes in sleep quality be expected to influence outcomes?) and in developing interventions designed to mitigate these problems (eg, the review will be able to identify which interventions are most effective).

Contributors The first author (AJS) had the idea for the proposed review and approached TLW and GR, who contributed to the design of the research. AJS drafted the protocol and TLW and GR provided detailed comments before submission. AJS is the identified guarantor of the review.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Colten HR, Altevogt BM. *Sleep disorders and sleep deprivation: an unmet public health problem*. Washington, DC: National Academies Press, 2006.
- Patel V, Flisher AJ, Hetrick S, *et al*. Mental health of young people: a global public-health challenge. *Lancet* 2007;369:1302–13.
- Robotham D. Sleep as a public health concern: insomnia and mental health. *J Public Ment Health* 2011;10:234–7.
- Prince M, Patel V, Saxena S, *et al*. No health without mental health. *The Lancet* 2007;370:859–77.
- Ford DE, Kamerow DB. Epidemiologic study of sleep disturbances and psychiatric disorders. An opportunity for prevention? *JAMA* 1989;262:1479–84.
- Baglioni C, Nanovska S, Regen W, *et al*. Sleep and mental disorders: a meta-analysis of polysomnographic research. *Psychol Bull* 2016;142:969–90.
- Benca RM, Obermeyer WH, Thisted RA, *et al*. Sleep and psychiatric disorders. A meta-analysis. *Arch Gen Psychiatry* 1992;49:651–68.
- Baglioni C, Battagliese G, Feige B, *et al*. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. *J Affect Disord* 2011;135:10–19.
- Gregory AM, Rijdsdijk FV, Lau JY, *et al*. The direction of longitudinal associations between sleep problems and depression symptoms: a study of twins aged 8 and 10 years. *Sleep* 2009;32:189–99.
- Breslau N, Roth T, Rosenthal L, *et al*. Sleep disturbance and psychiatric disorders: a longitudinal epidemiological study of young adults. *Biol Psychiatry* 1996;39:411–8.
- Harvey AG. Insomnia: symptom or diagnosis? *Clin Psychol Rev* 2001;21:1037–59.
- Soehner AM, Kaplan KA, Harvey AG. Insomnia comorbid to severe psychiatric illness. *Sleep Med Clin* 2013;8:361–71.
- Freeman D, Stahl D, McManus S, *et al*. Insomnia, worry, anxiety and depression as predictors of the occurrence and persistence of paranoid thinking. *Soc Psychiatry Psychiatr Epidemiol* 2012;47:1195–203.
- Cuijpers P, van Straten A, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. *J Behav Med* 2008;31:169–77.
- Wu JQ, Appleman ER, Salazar RD, *et al*. Cognitive behavioral therapy for insomnia comorbid with psychiatric and medical conditions: a meta-analysis. *JAMA Intern Med* 2015;175:1461–72.
- van Straten A, Cuijpers P. Self-help therapy for insomnia: a meta-analysis. *Sleep Med Rev* 2009;13:61–71.
- Koffel EA, Koffel JB, Gehrman PR. A meta-analysis of group cognitive behavioral therapy for insomnia. *Sleep Med Rev* 2015;19:6–16.
- Ho FY, Chung KF, Yeung WF, *et al*. Self-help cognitive-behavioral therapy for insomnia: a meta-analysis of randomized controlled trials. *Sleep Med Rev* 2015;19:17–28.
- Chan MS, Chung KF, Yung KP, *et al*. Sleep in schizophrenia: a systematic review and meta-analysis of polysomnographic findings in case-control studies. *Sleep Med Rev* 2017;32:69–84.
- Kobayashi I, Boarts JM, Delahanty DL. Polysomnographically measured sleep abnormalities in PTSD: a meta-analytic review. *Psychophysiology* 2007;44:660–9.
- Nota JA, Sharkey KM, Coles ME. Sleep, arousal, and circadian rhythms in adults with obsessive-compulsive disorder: a meta-analysis. *Neurosci Biobehav Rev* 2015;51:100–7.
- Reeve S, Sheaves B, Freeman D. The role of sleep dysfunction in the occurrence of delusions and hallucinations: a systematic review. *Clin Psychol Rev* 2015;42:96–115.
- Belleville G, Cousineau H, Levrier K, *et al*. Meta-analytic review of the impact of cognitive-behavior therapy for insomnia on concomitant anxiety. *Clin Psychol Rev* 2011;31:638–52.
- Lovato N, Gradisar M. A meta-analysis and model of the relationship between sleep and depression in adolescents: recommendations for future research and clinical practice. *Sleep Med Rev* 2014;18:521–9.
- Díaz-Román A, Perestelo-Pérez L, Buela-Casal G. Sleep in obsessive-compulsive disorder: a systematic review and meta-analysis. *Sleep Med* 2015;16:1049–55.
- Chouinard S, Poulin J, Stip E, *et al*. Sleep in untreated patients with schizophrenia: a meta-analysis. *Schizophr Bull* 2004;30:957–67.
- Cole MG, Dendukuri N. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. *Am J Psychiatry* 2003;160:1147–56.
- Zhai L, Zhang H, Zhang D. Sleep durations and depression among adults: a meta-analysis of prospective studies. *Depress Anxiety* 2015;32:664–70.
- Mauro R, Iovle U. Understanding L.O.V.E. (left out variables error): a method for estimating the effects of omitted variables. *Psychol Bull* 1990;108:314–29.

30. Zapf D, Dormann C, Frese M. Longitudinal studies in organizational stress research: a review of the literature with reference to methodological issues. *J Occup Health Psychol* 1996;1:145–69.
31. Duckworth AL, Tsukayama E, May H. Establishing causality using longitudinal hierarchical linear modeling: an illustration predicting achievement from self-control. *Soc Psychol Personal Sci* 2010;1:311–7.
32. Tang NK, Lereya ST, Boulton H, et al. Nonpharmacological treatments of insomnia for long-term painful conditions: a systematic review and meta-analysis of patient-reported outcomes in randomized controlled trials. *Sleep* 2015;38:1751–64.
33. Okajima I, Komada Y, INOUE Y. A meta-analysis on the treatment effectiveness of cognitive behavioral therapy for primary insomnia. *Sleep Biol Rhythms* 2011;9:24–34.
34. Hansen K, Höfling V, Kröner-Borowik T, et al. Efficacy of psychological interventions aiming to reduce chronic nightmares: a meta-analysis. *Clin Psychol Rev* 2013;33:146–55.
35. Gupta MA, Simpson FC, Lyons DC. The effect of treating obstructive sleep apnea with positive airway pressure on depression and other subjective symptoms: A systematic review and meta-analysis. *Sleep Med Rev* 2016;28:55–68.
36. Ho FY, Chan CS, Tang KN. Cognitive-behavioral therapy for sleep disturbances in treating posttraumatic stress disorder symptoms: a meta-analysis of randomized controlled trials. *Clin Psychol Rev* 2016;43:90–102.
37. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
38. Åkerstedt T, Hume K, Minors D, et al. The meaning of good sleep: a longitudinal study of polysomnography and subjective sleep quality. *J Sleep Res* 1994;3:152–8.
39. Harvey AG, Stinson K, Whitaker KL, et al. The subjective meaning of sleep quality: a comparison of individuals with and without insomnia. *Sleep* 2008;31:383–.
40. Libman E, Fichten C, Creti L, et al. Refreshing sleep and sleep continuity determine perceived sleep quality. *Sleep Disord* 2016;2016:1–10.
41. Lamarche LJ, De Koninck J. Sleep disturbance in adults with posttraumatic stress disorder: a review. *J Clin Psychiatry* 2007;68:1257–70.
42. Mellman TA, David D, Bustamante V, et al. Dreams in the acute aftermath of trauma and their relationship to PTSD. *J Trauma Stress* 2001;14:241–7.
43. Levin R, Fireman G. Nightmare prevalence, nightmare distress, and self-reported psychological disturbance. *Sleep* 2002;25:205–12.
44. Blake DD, Weathers FW, Nagy LM, et al. The development of a clinician-administered PTSD scale. *J Trauma Stress* 1995;8:75–90.
45. Buysse DJ, Reynolds CF, Monk TH, et al. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193–213.
46. Van de Water AT, Holmes A, Hurley DA. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography—a systematic review. *J Sleep Res* 2011;20:183–200.
47. Vaughn BV, Giallanza P. Technical review of polysomnography. *Chest* 2008;134:1310–9.
48. Beck AT, Steer RA, Carbin MG. Psychometric properties of the beck depression inventory: twenty-five years of evaluation. *Clin Psychol Rev* 1988;8:77–100.
49. Green CE, Freeman D, Kuipers E, et al. Measuring ideas of persecution and social reference: The Green et al. Paranoid thought scales (GPTS). *Psychological Medicine* 2008;38:101–11.
50. Johns LC, Cannon M, Singleton N, et al. Prevalence and correlates of self-reported psychotic symptoms in the British population. *Br J Psychiatry* 2004;185:298–305.
51. McGrath JA, Nestadt G, Liang KY, et al. Five latent factors underlying schizophrenia: analysis and relationship to illnesses in relatives. *Schizophr Bull* 2004;30:855–73.
52. Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry* 2007;4:28–37.
53. Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002;31:150–3.
54. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;62:1006–12.
55. Review manager (revman) [program]. 5.3 version. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration, 2014.
56. Lyons LC, Morris WA. Lyons-morris meta-analysis calculator 2017. <http://www.lyonsmorris.com/ma1/>
57. Bastien CH, Vallières A, Morin CM. Validation of the insomnia severity index as an outcome measure for insomnia research. *Sleep Med* 2001;2:297–307.
58. Scammacca N, Roberts G, Stuebing KK. Meta-analysis with complex research designs dealing with dependence from multiple measures and multiple group comparisons. *Rev Educ Res* 2014;84:328–64.
59. Cooper HM. *Integrating research: a guide for literature reviews*. London, England: Sage, 1986.
60. Cohen J. A power primer. *Psychol Bull* 1992;112:155–9.
61. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
62. Orwin RG. A fail-safe N for effect size in meta-analysis. *J Educ Stat* 1983;8:157–9.
63. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
64. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17:1–12.
65. Olivo SA, Macedo LG, Gadotti IC, et al. Scales to assess the quality of randomized controlled trials: a systematic review. *Phys Ther* 2008;88.
66. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383–94.
67. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
68. Lovibond PF, Lovibond SH. The structure of negative emotional states: Comparison of the depression anxiety stress scales (dass) with the beck depression and anxiety inventories. *Behaviour Research and Therapy* 1995;33:335–43.
69. Kenny DA, Kashy DA, Bolger N. *Data analysis in social psychology*. New York: Oxford University Press, 1998.
70. Barone JE. Comparing apples and oranges: a randomised prospective study. *BMJ* 2000;321:1569–70.