Investigating the Role and Treatment of Sleep in Relation to Psychotic Experiences

A thesis submitted to the University of Manchester for the degree of Doctor of Clinical Psychology (ClinPsyD) in the Faculty of Biology, Medicine and Health

2016

Gita Patel

School of Health Sciences
Division of Psychology and Mental Health
# List of Contents

List of Tables .................................................................................................................. 4  
List of Figures .................................................................................................................. 5  
Declaration ....................................................................................................................... 7  
Copyright Statement ...................................................................................................... 8  
Acknowledgements ......................................................................................................... 10  
The Author ...................................................................................................................... 11  
Abbreviations .................................................................................................................. 12  

**Paper 1**

Abstract ......................................................................................................................... 13  

1. Introduction .................................................................................................................. 15  
   1.1 Prevalence of Sleep Disturbance in Severe Mental Illness ................................. 15  
   1.2 Treatments for Sleep Disturbance ........................................................................ 17  
   1.3 Current Review: Aim and Scope ......................................................................... 22  

2. Method ......................................................................................................................... 23  
   2.1 Search Strategy ...................................................................................................... 23  
   2.2 Inclusion and Exclusion Criteria .......................................................................... 24  
   2.3 Screening procedure ............................................................................................ 24  
   2.4 Data Extraction & Synthesis ................................................................................. 27  
   2.5 Risk of bias/ Quality assessment .......................................................................... 27  

3. Results ......................................................................................................................... 28  
   3.1 Effect Size Calculations ...................................................................................... 36  
   3.2 Narrative Synthesis .............................................................................................. 40  

4. Discussion .................................................................................................................... 53  
   4.1 Clinical Implications ............................................................................................ 59  
   4.2 Future Directions ................................................................................................. 60  

5. Conclusions .................................................................................................................. 61  

6. References .................................................................................................................... 62
List of Tables

Paper 1

Table 1 Characteristics of Included Studies ......................................................................................................................... 30

Table 2 Post treatment effect size calculations for interventions used to treat sleep in severe mental illness ......................................................... 38

Table 3 Follow up effect size calculations for interventions used to treat sleep in severe mental illness ................................................................. 39

Paper 2

Table 1 Demographic characteristics of good sleepers and poor sleepers ................................................................. 88

Table 2 Comparisons between good sleepers and poor sleepers on questionnaires and computer tasks associated with psychotic-like experiences ......................................................... 90

Table 3 Correlation matrix detailing associations between sleep variables and performance on working memory and voice-detection task................................................................. 92

Paper 3

Table 1 Decibel readings for voices task across developer and test computer ........................................... 117
List of Figures

Paper 1

Figure 1 Prisma Diagram: Identification and screening of empirical studies included in systematic review ................................................................. 26

Paper 3

Figure 1 Basic Eligibility Criteria for Study Inclusion ........................................ 110
Figure 2 Recruitment Frequency Statistics ....................................................... 111
Figure 3 Reasons for Participant ineligibility at screening ............................... 115
Abstract

Sleep problems are a common complaint in clinical samples. Specifically, the high level of co-occurring sleep disturbances, such as insomnia and hypersomnia, amongst patients with schizophrenia and bipolar disorder suggests a relationship between poor sleep and psychopathology. Furthermore, recent evidence consistently indicates that sleep disturbance may be related to symptoms of psychosis such as hallucinations and delusions. The research described in this thesis attempts to explore the relationship between sleep in psychosis.

Paper 1 presents a systematic review entitled: Evaluating the Effectiveness of Sleep Treatments in Severe Mental Illness. The review identified fifteen controlled trials of pharmacological and non-pharmacological interventions for the treatment of sleep disturbance in patients with severe mental illness. The outcomes of the review indicated that sleep could be reliably improved in persons with schizophrenia and bipolar disorder. Significant improvements to sleep were observed for some pharmacological interventions as well as for non-pharmacological interventions.

Paper 2 describes an empirical study aimed at exploring the effects of poor sleep on the cognitive mechanisms underlying psychotic experiences. Using an independent groups design, two, well-defined, non-clinical groups comprising good sleepers and poor sleepers (those meeting criteria for insomnia disorder) were recruited to the study. The two groups were compared on a series of questionnaires and computer tasks designed to assess mechanisms underlying hallucinations and delusions, with a view to determine whether there was a difference in performance between groups that could be attributed to sleep. No significant differences were found between groups, although the study was underpowered. The findings are discussed in the context of sample characteristics and tests used to compare groups.

The final paper offers a critical reflection on the systematic review and empirical study, drawing together conclusions about the role and treatment of sleep in relation to psychotic experiences.
Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright Statement

1. The author of this dissertation (including any appendices and/or schedules to this dissertation) owns certain copyright or related rights in it (the “Copyright”) and she has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

2. Copies of this dissertation, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

3. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the dissertation, for example graphs and tables (“Reproductions”), which may be described in this dissertation, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

4. Further information on the conditions under which disclosure, publication and commercialisation of this dissertation, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (http://documents.manchester.ac.uk/display.aspx?DocID=24420), in any relevant Dissertation restriction declarations deposited in the University Library, The University Library’s regulations (see
http://www.library.manchester.ac.uk/about/regulations/) and in The University’s Guidance for the Presentation of Dissertations.
Acknowledgements

This thesis marks the end of my time at University and I am grateful for the continuous support that I have received from all of my friends and family not just over the past 3 years whilst completing the ClinPsyD, but over my ‘student career’ as a whole. Thanks especially to Jonathan Robson; for the tea, the hugs and the motivational pep-talks.

I would also like to thank my supervisors Gillian Haddock, Simon Kyle and Richard Brown for all of their guidance on the completion of this research project. Their expertise has been invaluable and I have learned so much under their direction.
The Author

Prior to embarking on the ClinPsyD the author worked as a Research Assistant across several forensic mental health settings. As part of these roles she engaged in a number of research and service evaluation projects. The author’s PhD was also completed in a High Secure Hospital and this took the form of a large, qualitative study.

BSc (Hons) Psychology  2006
MRes Psychology  2008
Doctor of Philosophy  2015
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCT</td>
<td>Randomised Controlled Trial</td>
</tr>
<tr>
<td>PSG</td>
<td>Polysomnography; objective method of sleep assessment that involves recording of physiological activity in the brain over a period of sleep. Polysomnographs can be used to monitor the brain, eye movements, heart rhythm and skeletal muscle activation during sleep. These data can be used to determine sleep parameters.</td>
</tr>
</tbody>
</table>

### Sleep Parameters:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOL</td>
<td>Sleep Onset Latency; time taken to fall asleep from lights out</td>
</tr>
<tr>
<td>WASO</td>
<td>Wake After Sleep Onset; time spent awake after sleep onset</td>
</tr>
<tr>
<td>TST</td>
<td>Total Sleep Time (duration)</td>
</tr>
<tr>
<td>SE</td>
<td>Sleep Efficiency; time asleep / time in bed expressed as a percentage</td>
</tr>
<tr>
<td>REM/NREM</td>
<td>Sleep usually consists of two phases: non-rapid eye movement sleep (NREM) and rapid eye movement (REM) sleep. NREM sleep consists of 3 stages (stage 1, stage 2, stage 3), each distinguished by different patterns of electrical brain activity.</td>
</tr>
<tr>
<td>SWS</td>
<td>Slow Wave Sleep; part of NREM (stage 3) represents a deeper sleep stage, often associated with memory consolidation.</td>
</tr>
</tbody>
</table>

### Common Subjective Assessments:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>Sleep Diary</td>
</tr>
<tr>
<td>PSQI</td>
<td>Pittsburgh Sleep Quality Index</td>
</tr>
<tr>
<td>ISI</td>
<td>Insomnia Severity Index</td>
</tr>
</tbody>
</table>

### Interventions:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBTI</td>
<td>Cognitive Behaviour Therapy for Insomnia</td>
</tr>
<tr>
<td>CBTI-BP</td>
<td>Cognitive Behaviour Therapy for Insomnia adapted for Bipolar Disorder</td>
</tr>
</tbody>
</table>
Paper 1

Evaluating the Effectiveness of Sleep Treatments in Severe Mental Illness (SMI): A Systematic Review

Prepared for submission to Clinical Psychology Review

Abstract

Sleep disturbances in schizophrenia and bipolar disorders are extremely common. Recent evidence suggests that sleep dysfunction may even predict or exacerbate psychotic experiences. As such, the treatment of sleep dysfunction in this context has become increasingly important. This paper presents a review of the effectiveness of sleep treatments in people with severe mental illness. Fifteen papers were identified through electronic searches of Medline, PsycInfo, Embase and PubMed databases. Studies were included if they sampled patients with schizophrenia or bipolar disorder, trialled a sleep intervention and reported on objective or subjective sleep outcomes. All studies were appraised for methodological quality. Effect sizes were calculated where sufficient data were presented, although a meta-analysis was not completed. Two studies trialled non-pharmacological interventions with the remaining trialling antipsychotics, hypnotics or melatonin receptor agonists. Thirteen studies were reviewed in the narrative synthesis. Findings indicated that sleep could be improved in patients with schizophrenia using melatonin, hypnotics, antipsychotics and CBTI. Antipsychotics and CBTI-BP could also be used to treat sleep disturbances in bipolar disorder. Future trials of sleep treatments in severe mental illness should be large, adequately powered, demonstrate rigorous and clear screening and report sleep outcomes using clinically validated measures.
Keywords: sleep disturbance; sleep treatment; insomnia; severe mental illness; schizophrenia; bipolar disorder.

Highlights

- No studies were found examining interventions specific to hypersomnia, circadian-rhythm sleep-wake disorder, or nightmare disorders, despite the prevalence of these conditions within schizophrenia and bipolar patient groups. Insomnia on the other hand was better-represented in the literature.
- CBTI and some pharmacological interventions demonstrated improvement to sleep in people with severe mental illness, although improvement was largely recorded through subjective self-report assessment.
- Large treatment effects were observed for the two pilot trials of CBTI, with improvements maintained at follow-up.
1. Introduction

1.1 Prevalence of Sleep Disturbance in Severe Mental Illness

There is a widely recognised association between sleep disturbance and psychotic experiences. Insomnia (difficulty initiating and/or maintaining sleep) is regarded as the most commonly occurring sleep disturbance amongst patients with schizophrenia, with prevalence rates ranging from 36% - 80% (e.g. Cohrs, 2008; Freeman, Pugh, Vorontova & Southgate, 2009; Palmese et al., 2011; Xiang et al., 2009). People with schizophrenia also have markedly impaired circadian rhythms ranging from highly irregular sleep patterns with normal melatonin production, to severe misalignment between the body’s circadian rhythm (i.e. melatonin production) and associated sleep-wake patterns (Wulff, Dijk, Middleton, Foster & Joyce, 2012).

Moreover, difficulties maintaining consistent sleep patterns can result in emotional distress and daytime fatigue amongst patients with schizophrenia (Waite, et al., 2015), which may exacerbate symptoms of psychosis. For example, a recent review by Reeve, Sheaves & Freeman (2015) found evidence to suggest that sleep dysfunction can predict the occurrence of psychotic experiences such as hallucinations and delusions, specifically insomnia with paranoia. In addition, a recent study by Mulligan, Haddock, Emsley, Neil & Kyle (2016) demonstrated more specifically that impairment in sleep continuity (e.g. sleep fragmentation and sleep efficiency) is implicated in the occurrence of hallucinations and delusions, and furthermore that these relationships are mediated by negative affect on awakening. These studies not only demonstrate the extent of sleep disturbance in schizophrenia but also highlight the complex relationship between sleep disruption and psychotic symptoms.
Sleep disturbance is also frequently observed in bipolar disorder. A reduced need for sleep is characteristic of individuals experiencing a manic episode, whilst both hypersomnia (excessive sleepiness) and insomnia are recognised as characteristic of a depressive episode (Kanady, Soehner & Harvey, 2015). Moreover, changes in sleep pattern are also evident during the inter-episode periods where up to 70% patients have reported insomnia (Harvey, 2008) and 25% reported hypersomnia (Kaplan, Gruber, Eidelman, Talbot & Harvey, 2011). The frequent reports of persistent sleep disturbance in bipolar disorder has led some to conclude that sleep dysregulation is a core mechanism in bipolar disorder (Harvey, Schmidt, Scarnà, Semler & Goodwin, 2005). People with bipolar disorder often report circadian rhythm dysfunction and indeed there is some evidence to suggest a genetic predisposition underpinning circadian rhythm instability in bipolar patients (Mansour, Monk & Nimgaonkar, 2005). Nevertheless, it is widely accepted that sleep disturbance, regardless of aetiology, contributes to affective dysregulation and is recognised as a common prodromal symptom indicative of relapse (Jackson, Cavanagh & Scott, 2003).

The cumulative effect of such sleeping difficulties are therefore considered to perpetuate mental health difficulties, by means of affecting how individuals are able to regulate their mood and engage in daily activities, which supports a bi-directional link between mental illness and sleep (Harvey, 2008). Despite high levels of co-occurrence, the treatment of sleep within the context of severe mental illness (SMI) has only recently become recognised as a primary symptom worthy of treatment in its own right, as opposed to a secondary symptom of psychosis or side-effect to antipsychotic medication (e.g. Wilson & Argyropoulos, 2012). Accordingly, Harvey (2009) argues that there should be impetus towards interventions for insomnia and other sleep disorders, for people with severe psychiatric disorders.
1.2 Treatments for Sleep Disturbance

The National Institute for Clinical Excellence (NICE) in the UK recommends both pharmacological and non-pharmacological interventions for the treatment of sleep disturbance and there are an increasing number of clinical trials examining the effectiveness of different interventions for sleep disturbance.

1.2.1 Treatments for sleep disturbance in non-psychiatric samples

A range of treatments are available for improving sleep in non-psychiatric samples and these include pharmacological as well as non-pharmacological treatments. Light therapy can be used to treat circadian-rhythm sleep-wake disorders (Dodson & Zee, 2010), although with varying rates of success (Morgenthaler et al., 2007). Morgenthaler et al., (2007) also identifies other treatment options which include melatonin (sleep hormone) and prescribed sleep schedules for circadian rhythm sleep disorders, depending on the variant of the condition.

Hypersomnia is thought to affect 0.01% – 0.02 % of the general population (Sowa, 2016). There have been few studies on the treatment of hypersomnia, but from the literature that is available there is evidence to suggest that Modafinil (a wakefulness promoting drug) appears to produce moderate improvements to sleep (Sowa, 2016). Sowa (2016) concluded that stimulants and other treatments such as melatonin may also lead to sleep improvements, although more trials are needed to reliably indicate effectiveness of these drugs.

Much more is known about the treatment of insomnia. Pharmacological treatments include prescription drugs which induce sedative effects e.g. benzodiazepines with hypnotic properties, or the use of exogenous sleep hormone (e.g. melatonin), although it is considered that this
only be used in adults aged 55 and over (Wilson et al., 2010). While these medications have been shown to have high levels of efficacy in treating insomnia symptoms (Holbrook, Crowther, Lotter, Cheng & King, 2000; Huedo-Medina, Kirsch, Middlemass, Klonizakis, & Siriwardena, 2012) the potential benefits are often outweighed by side-effects, such as daytime sedation and poor motor co-ordination, which also cause concern for driving accidents and injuries sustained from falls. Furthermore, there is evidence to suggest that some hypnotic drugs foster dependence (e.g. Benca, 2001; Kramer, 2000) and that progressive use of hypnotics increases tolerance thus reducing their effectiveness over time (NICE, 2015). As such, there is increasing evidence to suggest a move away from pharmacological interventions in favour of non-pharmacological approaches (Qaseem, Kansagara, Forciea, Cooke and Denberg, 2016).

In recently published clinical guidelines, Qaseem and colleagues (2016), on behalf of the American College of Physicians, recommend Cognitive-Behaviour Therapy for Insomnia (CBTI) for the initial treatment of insomnia in adults. CBTI is increasingly recognised in the treatment of insomnia (Morin & Benca, 2012) and is reported to be highly effective in non-clinical populations (e.g. Harvey et al., 2014; Morin, Culbert & Schwartz, 1994). CBTI components are cognitive (targeting beliefs and thoughts that increase arousal and interfere with sleep) and behavioural (to extinguish the association between efforts to sleep and actual sleep).

There is a growing body of research examining the use of non-pharmacological approaches to treating insomnia. For example, Morin et al. (1999, 2006) found that cognitive and behavioural approaches, relaxation, sleep-restriction interventions can produce reliable changes in several sleep parameters (to include sleep onset latency and sleep duration) for those with insomnia, and moreover that these improvements can be sustained over time (e.g. Morin et al., 2006). A recent review by Trauer, Qian, Doyle, Rajaratnam
& Cunnington (2015) also identified CBTI as an effective treatment for insomnia for adults who did not have comorbid psychiatric disorders.

1.2.2 Treatments for sleep disturbance in people diagnosed with schizophrenia and bipolar disorder

A systematic review by Cohrs (2008) assessed the influence of antipsychotic medications on sleep in schizophrenia and found that both first and second generation antipsychotic drugs were associated with improvements to sleep duration, sleep efficiency (SE) and continuity. Despite the sedative properties of antipsychotic drugs, they are not recommended as a direct treatment for sleep disturbance. While these pharmacological interventions can have a therapeutic impact (see Cohrs, 2008; Krystal, Goforth & Roth 2008; for reviews), there is potential for adverse effects to occur, which incidentally also affect sleep. Krystal, Goforth & Roth (2008) reported that these adverse effects could include insomnia/disturbed sleep (4-25%), weight-gain (0.3kg – 5.7kg over 10 weeks), which has also been associated with sleep apnoea, and Parkinsonian extra-pyramidal side-effects (3-65%)\(^1\). In the absence of a systematic synthesis of findings that weigh the consequences of pharmacological intervention against potential improvements, there continues to be debate as to the value of treating sleep disturbances in schizophrenia with antipsychotic drugs.

Furthermore, there are an increasing number of non-pharmacological treatments for sleep in schizophrenia (e.g. Freeman et al., 2015; Myers, Startup & Freeman, 2011) which offer potential for sleep improvement without the burden of side-effects associated with pharmacological approaches. Moreover, non-pharmacological approaches, such as cognitive

\(^1\) Data obtained from Cochrane reviews and meta-analyses (Allison et al., 1999; Bagnall et al., 2000; Sultana et al., 2000; Thornley et al., 2003; Srisurapanont et al., 2004; Duggan et al., 2005; El-Sayeh and Morganiti, 2006; Jayaram et al., 2006; Joy et al., 2006).
and behavioural therapies, are actually identified as the preferred sleep treatment option for patients with schizophrenia (Waters, Chiu, Janca, Atkinson & Ree, 2015). Although, again, there has been little systematic investigation of the effectiveness of non-pharmacological interventions for the treatment of common sleep problems in schizophrenia and, to date, no review that considers the effectiveness of both pharmacological and non-pharmacological interventions.

In bipolar disorder, pharmacological interventions used to manage symptoms of the disorder are inseparable from the treatment of sleep disturbance; typically, antipsychotic drugs are used to treat mood and regulate circadian rhythms in bipolar disorder. Despite the efficacy of pharmacological interventions however, relapse for bipolar disorder remains high (Gitlin, Swendsen, Heller & Hammen, 1995). Moreover, Harvey, Kaplan & Soehner (2015a) observe that even with good adherence to medication, some people with bipolar disorder will continue to remain symptomatic, even during the inter-episode period. Insomnia is one such residual symptom that may contribute to affective dysregulation and functional impairment (to daily activities, for example), thereby maintaining difficulties (Harvey, 2008). Whilst antipsychotic medication continues to function as a major treatment for the management of bipolar it is recognised that there is a need for improved integration of pharmacological interventions with non-pharmacological interventions, as this may not only improve adherence to pharmacological intervention (Geddes & Miklowitz, 2013) but may also lead to improved wellness for the client and have better economic outcomes (Miklowitz & Scott, 2009).

The range of evidence-based, non-pharmacological interventions in bipolar disorder include Cognitive Behaviour Therapy (CBT), Interpersonal and Social Rhythm Therapy (IPSRT; Frank, Swartz & Kupfer, 2000) and family focused approaches (e.g. Simoneau, Miklowitz, Richards, Saleem &
George, 1999). While these interventions include at least one treatment component aimed specifically at targeting sleep, sleep outcomes are rarely reported when these treatments are evaluated. This limits current understanding about the efficacy of non-pharmacological interventions for the treatment of sleep in bipolar disorder. Narrative reviews (e.g. Harvey, Kaplan & Soehner, 2015a) helpfully demonstrate a role for adjunctive non-pharmacological approaches in the treatment of sleep. Importantly, Harvey et al., (2015a) highlight a need for non-pharmacological interventions to draw upon advances in the application of CBTI, particularly given the increasing application and success of this approach in treating insomnia across different clinical groups (e.g. Trauer, Qian, Doyle, Rajaratnam & Cunnington, 2015; Wu et al., 2015). Whilst Harvey et al., (2015a) suggest that non-pharmacological interventions for sleep are advantageous, there remains a need for this to be empirically reviewed; though systematic appraisal of the available evidence.

Historically, there have been few studies on the effectiveness of treatments for sleep specifically in clinical samples, as interventions have generally been geared towards management of symptoms or distress and reporting on these outcomes accordingly. However, sleep is increasingly recognised as a symptom on its own merits and thus warrants direct intervention. The evidence base for the treatment of sleep is increasing and therefore it is timely to collectively review these findings.

There have been no systematic appraisals of controlled sleep treatment evaluations in the context of schizophrenia or bipolar disorder, which limits current knowledge about the range and efficacy of sleep treatments in SMI. A review of the effectiveness of pharmacological and non-pharmacological interventions designed to treat sleep in the context of SMI is therefore warranted.
1.3 Current Review: Aim and Scope

The current review aims to contribute to the growing body of research that recognises sleep as an important treatment target SMI. This will be achieved through a systematic review of empirical studies that have sampled individuals with schizophrenia and bipolar disorder, carefully administered sleep-related interventions and appropriately reported on sleep-related outcomes.

The review will focus specifically on sleep problems that are common in SMI. These include problems with the quality, duration, timing and maintenance of sleep as well as excessive daytime sleepiness. In this respect the review will include insomnia, hypersomnia and circadian-rhythm sleep-wake disorder, but will also include other sleep disorders that have been identified in samples of individuals with SMI e.g. nightmare disorders. Disorders which have a more organic nature such as sleep apnoea and narcolepsy will be excluded, as will those disorders that have less of an etiological link with SMI i.e. sleep-related movement disorders such as restless leg syndrome.

It is anticipated that the inclusion of differential sleep treatments (both pharmacological and non-pharmacological) will help to differentiate the effects of different sleep-focused treatments on different sleep disorders.

The review broadly aims to identify studies of sleep treatment in individuals with severe mental illness with a view to establishing what the evidence base is for the treatment of sleep in this context. The review questions are as follows:

1. What sleep treatments exist to treat common sleep disturbance (insomnia, hypersomnia, nightmares and circadian rhythm sleep disorders) within the context of severe mental illness?
2. What is the level and quality of evidence that supports the treatments of sleep within SMI?

3. How effective are sleep interventions within the context of SMI?

This review will also include a quality audit and calculation of effect sizes in order to determine the quality of evidence and magnitude of effects for the treatment of sleep in SMI.

2. Method

2.1 Search Strategy

The searches were performed in January 2016. Four electronic bibliographic databases were searched from 1980 - present: Medline (via OvidSP), PsycInfo (via OvidSP), Embase (via OvidSP) and PubMed (Via National Library of Medicine). The EU Clinical Trials Register and the UK Clinical Trials Gateway were also searched for relevant studies.

The search included both free-text terms and indexed keywords (e.g. MeSH). Search terms for *sleep* (Sleep OR ‘sleep disorder’ OR ‘sleep disturb*’ OR insomnia OR circadian OR nightmare OR hypersomnia) *psychosis* (Psychos* OR Psychotic OR Schizophreni* OR Schizoaffective OR Bipolar OR ‘bipolar disorder’ OR Delusion* OR Hallucination*) and *intervention* (‘Cognitive behaviour’* OR therapy* OR intervention OR treatment* OR medic* OR drug* OR hypnotic* OR melatonin OR chronotherapy) were combined with the AND Boolean operator to collate relevant papers. A full list of the search strategy is presented in Appendix A.

Grey literature such as conference proceedings, theses and book chapters were not included in the search.
2.2 Inclusion and Exclusion Criteria

Studies were included if they were written in English, appeared in peer-reviewed journals and were available in full-text format. Included studies sampled adults (18 years or older) diagnosed with schizophrenia or bipolar disorder according to the Diagnostic and Statistical Manual – 5th edition (DSM-V; American Psychiatric Association, 2013) or International Classification of Diseases – version 10 (ICD-10; World Health Organisation, 1992) classification systems to include schizophrenia, schizophreniform disorder, schizoaffective disorder and bipolar disorder (Bipolar I and Bipolar II). Studies were included if they reported on primary data derived from randomised and non-randomised controlled trials of pharmacological (e.g. melatonin) or non-pharmacological interventions (e.g. CBT) to treat sleep. Accordingly, the primary or co-primary outcome of the intervention of included studies was focused on a change to objectively or subjectively measured sleep. Studies were included if they reported on specific (pre and post) outcome measures relating to sleep.

Studies were excluded if they sampled schizotypal personality disorder, brief psychotic disorder, substance/medication induced psychotic disorders and psychotic-disorders due to another medical condition. Studies which sampled individuals with co-existing cognitive impairment or neurological disorder were also excluded. Mixed samples were also excluded if <50% had SMI. Studies published in grey literature, case studies, qualitative studies and studies that did not report on trial-based outcomes (e.g. uncontrolled studies) were excluded.

2.3 Screening procedure

The search strategy yielded 2839 articles. Following the removal of duplicates, one author (GP) independently screened the titles and abstracts
of all remaining articles (n=2369) and excluded all those that were definitely not relevant. The full text was reviewed for 48 articles. These papers were read and evaluated against the inclusion and exclusion criteria. Studies that did not clearly meet the inclusion criteria were reviewed and discussed with another member of the research team until agreement could be reached as to whether the study should be included. Reference lists of all included papers were also searched for additional references. Fifteen studies met inclusion criteria for the review.

Figure 1 presents a summary of the screening and study selection process and follows guidance from the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Moher, Liberati, Tetzlaff, Altman, 2009). The PRISMA statement (Moher et al., 2009) consists of a checklist and flow-diagram which are intended to improve the reporting of systematic reviews and meta-analyses. The PRISMA flow-diagram has been used in the current review to illustrate the 4 stages of the study selection process. In addition, one extra stage has been added to the diagram to acknowledge two studies that were not included in the narrative synthesis (following quality audit).
Records identified through database searching (n = 2838) 

Records after duplicates removed (n = 2369) 

Records excluded (n = 2321) 

Records identified through hand search of references (n = 1) 

Records screened (n = 2369) 

Full text articles excluded, with reasons (n=33) 

n=13 studies with inappropriate study design: 10 case series; 2 case-control; 1 cross sectional. 

n=5 did not meet population criterion (e.g. adolescents/mixed and poorly defined SMI sample). 

n=4 did not meet intervention criterion (i.e. did not trial an intervention). 

n=11 Did not meet outcome criterion i.e. did not report on sleep outcome variables. 

Full text articles assessed for eligibility (n = 48) 

Studies meeting inclusion criteria (n=15) 

Excluded due to unreliable data (n=2) 

Studies included in qualitative synthesis (n=13) 

Figure 1 Prisma Diagram: Identification and screening of empirical studies included in systematic review
2.4 Data Extraction & Synthesis

A standardised data collection proforma was used to extract relevant data from the included studies for the assessment of study quality and evidence synthesis. Due to the heterogeneity across primary studies, a narrative synthesis methodology was employed to systematically extract and present findings. The guidance on narrative synthesis produced by Popay et al., (2006) was used to organise and synthesise findings in relation to the effectiveness of different interventions used to treat sleep in people with SMI.

Data relating to study characteristics, intervention (including effectiveness) and outcome measurement were tabulated and are presented across Tables 1, 2 and 3 as part of the preliminary synthesis.

2.5 Risk of bias/ Quality assessment

The quality assessment checklist developed by Downs & Black (1998) was used to assess the quality of studies included in the review. This has been regarded as a useful tool for evaluating randomised and non-randomised studies (Deeks et al., 2003). The Downs and Black quality assessment tool has 27 items and prompts review of quality across 5 domains: reporting, external validity, bias (internal validity), confounding (internal validity) and power. As such, the measure facilitates recognition of a study’s relative strengths and weaknesses. The reporting domain is designed to ascertain whether the study reporting is sufficient to making an unbiased assessment. One item on the checklist relates to power and following a statistical calculation, aims to identify whether negative findings could be due to chance. Up to 5 points could be awarded for this item, however, given lack of clarity around how calculations are assigned a corresponding score a decision was made to amend this item. The power
item was therefore simplified to yield a binary output instead; studies that completed a power calculation to estimate sample size were scored 1 and studies that did not scored 0. Scores on the checklist (with amendment to item 27) could range from 0 to 28. Quality ratings were completed by author GP and 25% of papers were also reviewed and rated by a second independent rater. Agreement between raters was moderate ($k=0.68, p<0.05$). Discrepancies were reviewed through consensus discussion and agreement between the raters.

3. Results

Fifteen studies met criteria for inclusion in the review. Twelve studies were RCTs (Goder et al., 2008; Kim et al., 2014; Kluge et al., 2014; Kumar et al., 2007; Luthringer et al., 2007; McElroy et al., 2010; Norris et al., 2013; Tek et al., 2014, Wamsley et al., 2013; Yamashita et al., 2004), of which two were pilot RCTs (Freeman et al., 2015; Harvey et al., 2015b). Two studies employed a randomised crossover trial design (Leibenluft et al., 1997; Shamir et al., 2000) and one used a non-randomised controlled trial methodology (Dursun et al., 1999). Eight medication RCTs were placebo-controlled (Goder et al., 2008, Kumar et al., 2007; Leibenluft et al., 1997; McElroy et al., 2010; Norris et al., 2013; Shamir et al., 2000; Tek et al., 2014; Wamsley et al., 2013).

Twelve studies specified changes in sleep as their primary outcome of the trialled intervention (Dursun et al., 1999; Freeman et al., 2015; Harvey et al., 2015b; Kim et al., 2014; Kluge et al., 2014; Kumar et al., 2007; Luthringer et al., 2007; McElroy et al., 2010; Norris et al., 2013; Shamir et al., 2000; Tek et al., 2014, Yamashita et al., 2004).

The majority of studies (13) trialled pharmacological interventions and just 2 studies (Freeman et al., 2015; Harvey et al., 2015b) trialled non-pharmacological interventions, both CBTI. Five studies sampled individuals
with bipolar disorder (Harvey et al., 2015b; Kim et al., 2014; Leibenluft et al., 1997; McElroy et al., 2010; Norris et al., 2013) and the remaining sampled individuals with schizophrenia related disorders and included both inpatients and outpatient samples.

Characteristics of included studies are presented in Table 1. Data presented in Table 1 are grouped according to schizophrenia and bipolar patient studies and within each group, studies are organised by intervention type, clustering similar pharmacological interventions together to facilitate comparison.
# Characteristics of Included Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Sample characteristics</th>
<th>Sleep Intervention (duration and dosage information)</th>
<th>Comparator</th>
<th>Sleep assessments</th>
<th>Quality Rating Total /28</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Studies examining sleep in schizophrenia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Subjective: ISI, PSQI, SD.</td>
<td></td>
</tr>
<tr>
<td>Goder et al., (2008)</td>
<td>RCT</td>
<td>N=26 patients diagnosed with schizophrenia: n=25 inpatients; 27% female.</td>
<td>Pharmaceutical: 10mg Olanzapine administered at 10pm on treatment night.</td>
<td>Olanzapine compared to placebo.</td>
<td>Objective: PSG Subjective: PSQI and analogue scale (0-9).</td>
<td>18</td>
</tr>
<tr>
<td>Kluge et al., (2014)</td>
<td>RCT</td>
<td>N=30 patients with schizophrenia/ schizoaffective disorder/ schizophreniform disorder. 60% female.</td>
<td>Pharmacological: Olanzapine (5-25mg/day) or Clozapine (100 – 400 mg/day).</td>
<td>Olanzapine vs. Clozapine</td>
<td>Objective: PSG</td>
<td>20</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>---------------</td>
<td>-----</td>
</tr>
<tr>
<td>Yamashita et al., (2004)</td>
<td>RCT</td>
<td>N=92 schizophrenia inpatients ($n=29$ disorganised; $n=11$ paranoid; $n=52$ undifferentiated)</td>
<td>Pharmaceutical: olanzapine (2.5-20mg per day)/ risperidone (1.0-12.0mg per day)/ perospirone (4.0 – 48.0mg per day)/ quetiapine (50.0-750mg per day).</td>
<td>Typical vs. atypical antipsychotic drugs (1 of 4)</td>
<td>Subjective: PSQI</td>
<td>18</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>N</td>
<td>Diagnosis</td>
<td>Average age</td>
<td>Sex distribution</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------</td>
<td>----</td>
<td>-----------------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Kumar et al., (2007)</td>
<td>RCT</td>
<td>40</td>
<td>Outpatients with paranoid schizophrenia</td>
<td>37 years</td>
<td>32% female</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharmacological:</td>
<td></td>
<td></td>
<td>Melatonin (15 days). 1 capsule (3mg melatonin) per night for nights 1 and 2. Thereafter flexible dosing permitted every other night (up to 4 capsules/12mg).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjective:</td>
<td></td>
<td></td>
<td>Melatonin vs. placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shamir et al., (2000)</td>
<td>Randomised crossover trial</td>
<td>19</td>
<td>Outpatients with schizophrenia (n=9 paranoid; n=5 disorganised; n=5 schizoaffective disorder). Average age 42 years; 36% female.</td>
<td>42 years</td>
<td>36% female</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharmaceutical:</td>
<td></td>
<td></td>
<td>Melatonin (3 weeks). 2mg taken once daily, 2 hours before bedtime for 3 weeks. 1 week washout between treatment periods.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective:</td>
<td></td>
<td></td>
<td>Melatonin vs. placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tek et al., (2014)</td>
<td>RCT</td>
<td>39</td>
<td>Patients with schizophrenia (n=20 schizophrenia; n=19 schizoaffective disorder). Average age 46 years; 50% female.</td>
<td>46 years</td>
<td>50% female</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharmaceutical:</td>
<td></td>
<td></td>
<td>Eszopiclone (8 weeks). 2mg each night before bedtime for week 1, 3mg per night from week 2-8.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjective:</td>
<td></td>
<td></td>
<td>Eszopiclone vs. placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISI and SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>N</td>
<td>Participants</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Objective</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>---</td>
<td>--------------</td>
<td>--------------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>Wamsley et al., (2013)</td>
<td>RCT</td>
<td>N=21 outpatients with schizophrenia. Average age 34 years, 19% female.</td>
<td>Pharmaceutical: eszopiclone (2 nights). 3mg 30minutes prior to bed time.</td>
<td>Eszopiclone vs. placebo</td>
<td>Objective: PSG</td>
<td>22</td>
</tr>
<tr>
<td>Dursun et al., (1999)</td>
<td>Non-randomised controlled trial</td>
<td>N=16 patients diagnosed with schizophrenia. Experimental/ control groups matched for gender.</td>
<td>Pharmaceutical: Risperidone (dosage and duration unclear)</td>
<td>Risperidone compared to typical antipsychotic drugs (APDs) and control group (non-clinical).</td>
<td>Objective: Actigraphy Subjective: Analogue scale to rate sleep quality and morning sleepiness (0-10).</td>
<td>11</td>
</tr>
<tr>
<td>Harvey et al., (2015b)</td>
<td>RCT</td>
<td>N=58 patients with Bipolar I disorder. 62% female. Average age 36 years.</td>
<td>Psychological: CBTI-BP (8 weeks). Psychoeducation (PE)</td>
<td></td>
<td>Subjective: ISI, DSISD, SD, PSQI, PROMISE-SD</td>
<td>24</td>
</tr>
<tr>
<td>Kim et al., (2014)</td>
<td>RCT (open-label trial)</td>
<td>N=29 patients with Bipolar I or Bipolar II (89.7%). Average age 36 years; 55% female.</td>
<td>Pharmacological: Quetiapine XR (8 weeks). Target dose of 300mg/day, by day 7 of trial.</td>
<td>Quetiapine vs. Lithium</td>
<td>Objective: Actigraphy Subjective: PSQI</td>
<td>19</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Comparison</td>
<td>Subjective Measures</td>
<td>Objective Measures</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
<td>-------------------------------------</td>
<td>---------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Leibenluft et al., (1997)</td>
<td>Randomised cross-over trial. Double-blind, placebo-controlled.</td>
<td>Five female patients with rapid-cycling bipolar disorder. Mean age 47.2 years (SD 3.8).</td>
<td>Pharmacological: Melatonin (12 weeks). 10mg/day administered at 22.00 each night in addition to usual medication</td>
<td>Melatonin vs. placebo</td>
<td>Melatonin concentration</td>
<td>12</td>
</tr>
</tbody>
</table>

Note. Abbreviations for subjective and objective measures: Objective: PSG = Polysomnography. Subjective Sleep: SD = Sleep Diary, PSQI = Pittsburgh Sleep Quality Index (Buysse, Reynolds, Monk, Berman & Kupfer, 1989); ISI = Insomnia Severity Index (Morin, 1993); LSEQ = Leeds Sleep Evaluation Questionnaire (Parrott & Hindmarch, 1980); PIRS = Pittsburgh Insomnia Rating Scale (Moul, Pilkonis, Miewald, Carey & Buysse, 2002); DSISD = Duke Structured Clinical Interview for Sleep Disorders (Edinger, et al., 2004); PROMIS-SD = Patient-Reported Outcomes Measurement Information System – Sleep Disturbance (Buysse et al., 2010).
The total quality rating for each study is provided in Table 1 but further details on how this score was derived can be found by consulting the full scored checklist presented in Appendix B. The studies varied in quality with ratings ranging from 11 to 27 out of a possible score of 28.

During the quality appraisal process it became apparent that there was a lack of clarity in reporting for two studies. In the study by Dursun et al., (1999) there were some errors in reporting relating to risperidone dosage and duration (e.g. risperidone dosage was first specified as 9.5mg/day for 42.5 weeks and later in the article it was reported that this group received 42.5mg/day for 9.5 weeks). As such it was difficult to decipher the exact details of the intervention. Consequently, outcome data could not be reliably synthesised for inclusion in this review.

The study by Leibenluft et al., (1997) which trialled melatonin in bipolar disorder was also excluded from the narrative synthesis. Leibenluft et al., sampled just 5 patients and there was a lack of clarity in the reporting of these data, specifically around which patients completed the trial; some patients, it is not clear how many, were withdrawn/ did not complete due to adverse effects. In addition, the outcome data that are reported are limited to just three patients, yet it is unclear whether these data relate to the patients who either fully or partially completed the study. This is a significant limitation given the very small sample. As the study end-points could not be determined, a decision was made to exclude this paper from narrative synthesis.

The two papers that were excluded also had the lowest quality ratings with Dursun et al., (1999) scoring 11/28 and Leibenluft et al.,
(1997) scoring 12/28. However, it should be noted that both studies were excluded on the basis of their presenting unreliable information, as opposed to low quality scores; this review did not specify a minimum quality score for inclusion. Therefore, another study, which had only a marginally higher score, was maintained for narrative synthesis because it presented data that could be reliably extracted for analysis.

Following the removal of these two studies, 13 papers remained for inclusion within the narrative synthesis.

3.1 Effect Size Calculations

There was much variation in measurement of sleep across the set of papers reviewed; each study often reported on both subjective (e.g. sleep diaries, self-report measures and rating scales) and objective measures of sleep (e.g. polysomnography, actigraphy) and consequently there were several potential dependent variables. Furthermore, the methods of assessment varied across the set of papers reviewed; for example, within the subjective domain, there were 5 different questionnaires used to assess sleep, in addition to sleep diaries and simple analogue rating scales. Given the limited consistency in measurement it was not possible to calculate a meaningful effect size that was common to all studies. Accordingly, effect size calculations were limited to only those papers that reported on global measures of sleep disturbance (PSQI - Pittsburgh Sleep Quality Index; Buysee, Reynolds, Monk, Berman & Kupfer, 1989 or ISI – Insomnia Severity Index; Morin, 1993).

Eight studies reported on global measures of sleep disturbance. These were Freeman et al., (2015), Harvey et al., (2015), Kim et al., (2014), McElroy et al (2010), Goder et al., (2008), Tek et al.,
(2014), Norris et al., (2013), and Yamashita et al., (2004). However, effect size calculations could not be calculated for 4 studies (Goder et al., 2008; Norris et al., 2013; Tek et al., 2014; and Yamashita et al., 2004) as sufficient data were not presented to facilitate an effect size calculation. Therefore effect sizes were only calculated for 4 studies: Freeman et al., (2015), Harvey et al., (2015), Kim et al., (2014) and McElroy et al (2010).

The standardised mean difference (SMD), using Cohen’s $d$ was calculated for these studies and these data are provided in Table 2 (post-intervention effect size) and Table 3 (follow-up intervention effect size).
Table 2

*Post treatment effect size calculations for interventions used to treat sleep in severe mental illness*

<table>
<thead>
<tr>
<th>Study</th>
<th>Assessment Point</th>
<th>Primary sleep variable</th>
<th>Intervention Group (post-treatment)</th>
<th>Control Group (post-treatment)</th>
<th>Effect size (SMD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeman et al., (2015)</td>
<td>12 weeks</td>
<td>ISI total</td>
<td>N 22 M 9.3 SD 5.5</td>
<td>N 25 M 15.4 SD 5.4</td>
<td>-1.12</td>
</tr>
<tr>
<td>Harvey et al., (2015b)</td>
<td>8 weeks</td>
<td>ISI total</td>
<td>N 30 M 6.45 SD 5.49</td>
<td>N 28 M 13.90 SD 5.32</td>
<td>-1.38</td>
</tr>
<tr>
<td>Kim et al., (2014)</td>
<td>8 weeks</td>
<td>PSQI total</td>
<td>N 12 M 4.8 SD 3.6</td>
<td>N 17 M 6.1 SD 4.2</td>
<td>-0.33</td>
</tr>
<tr>
<td>McElroy et al., (2010)</td>
<td>8 weeks</td>
<td>PSQI total</td>
<td>N 6 M 8.3 SD 3.4</td>
<td>N 7 M 7.0 SD 4.2</td>
<td>0.34</td>
</tr>
</tbody>
</table>

*Note: Abbreviations- ISI - Insomnia Severity Index; PSQI – Pittsburgh Sleep Quality Index.*
Table 3

Follow-up effect size calculations for interventions used to treat sleep in severe mental illness

<table>
<thead>
<tr>
<th>Study</th>
<th>Follow-up Assessment Point</th>
<th>Primary sleep variable</th>
<th>Intervention Group (follow-up)</th>
<th>Control Group (follow-up)</th>
<th>Effect size SMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeman et al., (2015)</td>
<td>24 weeks</td>
<td>ISI</td>
<td>23 11.0 5.6</td>
<td>25 15.0 5.7</td>
<td>-0.71</td>
</tr>
<tr>
<td>Harvey et al., (2015b)</td>
<td>6 months</td>
<td>ISI</td>
<td>30 6.82 6.11</td>
<td>28 12.63 5.21</td>
<td>-1.02</td>
</tr>
</tbody>
</table>

Note: Abbreviations - ISI - Insomnia Severity Index; PSQI – Pittsburgh Sleep Quality Index.
A meta-analysis could not be completed due to the heterogeneity of included studies; variation in quality, different interventions (e.g. different medications trialled within the pharmacological interventions) and inconsistencies in the way in which sleep outcomes were recorded and reported precluded a statistical synthesis of the data.

3.2 Narrative Synthesis

The narrative synthesis presents data across two sections to reflect the two clinical populations included in this review: patients with schizophrenia and patients with bipolar disorder. These data are then summarised in the context of successful and unsuccessful treatments for sleep. Extracted data include details of sleep interventions, sleep assessments and sleep outcomes for each clinical group. Consideration is also given to the severity of sleep impairment in the clinical samples tested, so that the outcomes of interventions can be discussed in this context.

3.2.1 Studies examining sleep treatments in schizophrenia

Non-pharmacological interventions

The pilot trial of CBTI by Freeman & Colleagues (2015) was the only non-pharmacological intervention for the treatment of sleep in schizophrenia that was included in this review. The CBTI intervention² included psychoeducation, stimulus control therapy (e.g. setting of appropriate and regular sleep times, reducing sleep in the daytime) and the establishment of appropriate daytime activity and circadian rhythms (e.g. obtaining natural light in the morning). Specific cognitive techniques were also used to address unhelpful thoughts about sleep and to challenge particular problems relating to hallucinations and/or delusions that interfered with sleep. Ultimately, the

² manual providing detailed description of the approach can be found by consulting Waite et al., (2015).
specific components involved in the intervention were guided by the client’s needs and goals.

Freeman et al., (2015) used both objective (actigraphy) and subjective methods of assessment, although the actigraphy data was less complete, in part because many patients did not agree to wearing wrist actigraphy. The actigraphy data are reported and have small - moderate effect sizes. Similarly, data from the sleep diary that patients were asked to keep was not fully complete and again there were small – moderate effect sizes for improvements in these subjectively reported sleep variables.

Freeman et al., (2015) used clinically validated questionnaire measures to record changes in sleep. The primary outcome measure was a change in sleep and the results indicated reduction in ISI scores (indicating a reduction in insomnia severity) with large effect sizes at weeks 12 and moderate-large effect sizes at week 24 (see Tables 2 and 3). Although, the Freeman et al., study sampled n=50 patients and thus was only powered to detect only the largest of effect sizes.

The PSQI was also used to record changes in sleep quality and reductions in scores (indicating improved sleep quality) were also found following CBTI treatment. It should be noted that all patients recruited to this study had clinical levels of insomnia at baseline (as a pre-requisite for inclusion). A limitation of this pilot trial is the absence of an active control arm against which relative improvements to sleep can be considered, given strong placebo effects in sleep trials (Perlis, McCall, Jungquist, Pigeon & Matteson, 2005).

**Pharmacological Interventions**

Eight other studies trialled a range of pharmacological interventions in patients with schizophrenia, including melatonin (Kumar et al., 2007; and Shamir et al., 2000), eszopiclone (hypnotic) (Tek et al., 2014; Wamsley et al.,
Melatonin

Exogenous melatonin was trialled in two samples of patients with schizophrenia who complained of sleep disturbances and both trials of melatonin found significant improvement to sleep parameters following treatment. Kumar et al., (2007) sampled patients who complained of insomnia, self-reported sleep onset latency (SOL) of >30 minutes and reported that these difficulties persisted over 2 weeks and were distressing. Kumar et al., found that melatonin (3mg) was advantageous over placebo in contributing to improved sleep across 3 domains: sleep duration, night-time awakenings and sleep-onset latency (SOL). At the end of treatment, patients reported increased sleep duration and fewer night-time awakenings. While SOL had also decreased, this was not statistically significant.

Shamir et al., (2000) sampled patients with a formal diagnosis of insomnia disorder. Shamir and colleagues found that over 3 weeks of treatment with melatonin (2mg) patients in the treatment group demonstrated increased sleep efficiency (SE) compared to patients in the placebo group. Other objectively recorded sleep parameters were collated using actigraphy: SOL, sleep duration, wake after sleep onset (WASO), fragmentation index, and number of night-time awakenings, although no significant differences between groups emerged across these domains. Furthermore, SE was markedly improved for patients whose SE during placebo arm was below the group average. In this group of individuals with relatively low SE, treatment with melatonin appeared to contribute to reductions in SOL and increases in total sleep time (TST), although these improvements failed to reach significance.

A relative strength of both melatonin studies is that they were placebo-controlled. Both studies noted improvements to sleep measured
subjectively (Kumar et al., 2007) and objectively (Shamir et al., 2000), although the studies varied in quality with Kumar et al., (2007) scoring 21/28 compared to Shamir et al., (2000) scoring 13/28, given limited reporting and external validity controls in their study. In addition, it should also be noted that the findings from Shamir et al., are based on a relatively small sample of \( n=19 \) patients (\( n=40 \) recruited by Kumar et al.,).

**Hypnotics**

Tek et al., (2014) trialled eszopiclone (2mg) with a view to assessing its effects on reducing insomnia severity. Schizophrenia patients with existing sleep complaints (self-reported sleep difficulties twice per week in the past month, ISI score of \( \geq 10 \)) were recruited to the study. Patients were asked to keep a sleep diary each day, which was supplemented with a sleep interview each morning conducted by a blind assessor between weeks 1 and 2 and in the final week of drug administration. Sleep diary items included total sleep time (TST), wake after sleep onset (WASO), sleep onset latency (SOL), awakenings and also asked patients to rate alertness, concentration, physical wellbeing and functioning. Tek et al., (2014) found significantly reduced scores for insomnia severity for the eszopiclone group, compared with placebo. Sleep diary data relating to TST, WASO and SOL also improved in both groups across the trial period, although the difference between groups was not statistically significant.

Items from the sleep diary which included: *general feeling upon rising* and *concentration problems* were significantly improved in the eszopiclone group compared with the placebo group. The study by Tek et al., was high in methodological quality (scoring 27), however it should also be noted that the study was underpowered with a sample of \( n=39 \) recruited when \( n=80 \) were required for sufficient power.

Wamsley et al., (2013) also administered eszopiclone (3mg) with a view to determining its impact on sleep, although it should be noted that this
was not a therapeutic intervention trial of eszopiclone. Wamsley and colleagues were interested in assessing how the drug contributed to changes in sleep architecture and how this in turn affected memory performance. Sleep was assessed objectively using polysomnography. Wamsley and colleagues did not find any significant differences between placebo and eszopiclone groups for any measures of sleep quality or architecture. These findings are in contrast to Tek et al., (2014) although it should be noted that Wamsley et al., had a smaller sample than Tek et al., \( n=21 \) vs. \( n=39 \) and administered eszopiclone over a much shorter period of 2 nights, compared to Tek and colleagues’ 8-week trial. It should also be noted that the sleep characteristics of participants in each sample were different; with Wamsley et al., specifically excluding patients with diagnosed sleep disorders and those who had been treated with sleep medication. These may provide some account for disparate findings between the two studies of eszopiclone.

**Antipsychotics**

Luthringer et al., (2007) trialled paliperidone (9mg) in a sample of patients with schizophrenia and schizophrenia related insomnia. Although no formal assessment was used to screen severity of sleep disturbance, eligible participants were included if they complained of 1.5 hours of wakefulness per 8 hours in bed. Furthermore, baseline assessments indicated that patients in the trial (both groups) had severe sleep continuity disturbances, including prolonged SOL and WASO. Patients were assessed using objective PSG at baseline and at the end of the treatment period. Compared to patients in the placebo arm, patients treated with paliperidone had improved sleep architecture; paliperidone significantly increased both the duration and percentage of stage 2 sleep and the duration of REM sleep. Improvements to sleep quality were also reported by patients; both paliperidone and placebo groups similarly demonstrated improved scores at
the end of treatment, although they did not differ significantly from each other at end point. In addition, SOL was reduced in the paliperidone group compared to the placebo group, although this failed to reach significance. Luthringer et al., (2007) had a relatively high quality rating (score of 25) indicating a robust methodology underpinning these findings.

Yamashita et al., (2004) aimed to examine changes in sleep quality following change from typical to atypical antipsychotic drugs. Varying dosages within specified ranges were permitted for this trial (see Table 1 for dosage information) and this was determined by individual clinical status, adjusted by a psychiatrist. Mean dose levels at endpoint were: 16.5mg/day for olanzapine; 37.3mg/ day for perospirone; 432.5mg/ day for quetiapine and 7.4mg/day for risperidone. All patients included in the study had medication switched from typical to atypical antipsychotics and there was no control group for comparison. Overall, significant improvements were observed between baseline and the end of treatment for three of 4 trialled drugs (olanzapine, risperidone and quetiapine) across the following domains of the PSQI: sleep quality, SOL, SE, sleep disturbances and daytime dysfunction when compared to baseline. Mean PSQI total at baseline for the whole sample was 8.58 (SD 3.80) and at endpoint had reduced to 7.20 (SD 4.29) which was significant. Perospirone was the only drug treatment that did not result in significant change in sleep outcomes across any of the PSQI domains. Yamashita et al., (2004) also noted poor sleep quality at baseline could predict improvement to subjective sleep quality.

Olanzapine was trialled in a further two separate studies. Goder et al., (2008) measured the effects of olanzapine (10mg) using a combination of objective and subjective sleep assessments. Similar to Wamsley et al., (2013) described earlier, Goder and colleagues were interested in changes to sleep architecture and the impact this had on overnight memory consolidation. As
such, there also was a short period of administration of just 1 night in this study.

In the study by Goder et al., (2008) sleep complaints/disorders were not specified as part of the inclusion or exclusion criteria although baseline PSQI data are presented which indicate that the Mean PSQI score for the olanzapine group was 9.9 (SD 3.5) and 8.3 (SD 3.5) for the placebo group and these means were not significantly different at baseline. Sleep was assessed using both subjective and objective methods (PSG). The outcomes of the objective assessments indicated that olanzapine, relative to placebo, contributed to physiological change (significant increases in slow wave sleep and decreases in the amount of REM sleep). However, olanzapine failed to demonstrate significant change in sleep quality as assessed by the PSQI between the baseline and treatment night. It should be noted that the PSQI should be rated according to sleep experiences over a period of 1 month, therefore it is likely that this measure was not adequate to detect change over 1 night.

Kluge et al., (2014) also trialled olanzapine relative to clozapine treatment (no control condition). The dose was restricted during the first two weeks of intervention. The mean modal dose at endpoint was 266 mg/day of clozapine and 21.2 mg/day of olanzapine. PSG methods were used to objectively assess sleep. Kluge et al., found a significant increase in TST, SE and a decrease in SOL following olanzapine treatment. Improvements were observed in both groups, occurred early in treatment and were sustained at follow-up (6 weeks), although olanzapine was observed to improve sleep to a greater extent compared to clozapine treatment. Despite similar effects on global sleep parameters olanzapine and clozapine had differing effects on sleep architecture. Clozapine resulted in a much stronger increase of light, stage 2 sleep and olanzapine resulted in an increase in SWS, which was also observed by Goder et al., (2008). Both studies trialling olanzapine have
similar, moderate quality ratings; Kluge et al., scoring 20 and Goder et al., scoring 18 and these findings should be considered in this context.

From the studies examining sleep treatments in schizophrenia there is some preliminary evidence that CBTI is effective as a non-pharmacological treatment and that hypnotic eszopiclone, melatonin and some atypical antipsychotic drugs (i.e. olanzapine and quetiapine) may be associated with improvement to sleep. Although, these data have been obtained from studies that range in methodological quality (e.g. no placebo/control groups for Freeman et al., 2015; Kluge et al., 2014; Yamashita et al., 2004) and with generally small sample sizes (ranging from $n=19$ – Shamir et al., 2000; $n=21$ Wamsley et al., 2013) and the findings should be considered in this context.

3.2.2 Studies examining sleep treatments in bipolar disorder

Four studies that sampled patients with bipolar disorder are reviewed here (Harvey et al., 2015b; Kim et al., 2014; McElroy et al., 2010; Norris et al., 2013).

Non-Pharmacological

Cognitive Behaviour Therapy for Insomnia for Bipolar Disorder (CBTI-BP), is an adapted form of CBT designed to target unique features of sleep disturbance in bipolar disorder. CBTI-BP was trialled by Harvey et al., (2015b) in a sample of patients with bipolar disorder who also met criteria for insomnia disorder. Some patients included in the study also had difficulties with delayed sleep phase and hypersomnia, which necessitated inclusion of elements of chronotherapy and Interpersonal Social Rhythm Therapy (IPSRT) as part of the CBTI-BP intervention.

CBTI-BP components were delivered across three modules: behavioural (including stimulus control and sleep restriction), cognitive (challenging unhelpful beliefs around sleep) and relapse prevention. All modules were compulsory although the concentration on any one module was allowed to vary in line with individual case formulation. The control
group received psychoeducation (PE). Both groups received the same number of sessions (8) and the PE group controlled for therapeutic attention and expectation of improvement.

Sleep was assessed using clinically reliable measures which included the Insomnia Severity Index (ISI; Morin, 1993) and the Duke Structured Clinical Interview for Sleep Disorders (DSISD; Edinger, et al., 2004) as primary sleep outcome variables, and the PSQI as a secondary sleep outcome variable. Patients were also asked to keep sleep diaries to record SOL, WASO, TST, time in bed and terminal wakefulness. These data were also used to calculate SE and total wake time.

The outcomes of the trial indicated that both groups evidenced reduction in insomnia severity, although the CBTI-BP group improved to a greater extent than the control group at post-treatment and at follow-up. When compared with psychoeducation (control group), CBTI-BP was associated with significantly improved ISI scores post-treatment and at follow up with large effect sizes (see Tables 2 and 3). There were significantly more participants who no longer met diagnostic criteria for insomnia (according to the DSISD) in the CBTI-BP group post-treatment, although this improvement did not reach statistical significance at follow-up. The sleep diary data indicated similar results although with smaller effect sizes.

Pharmacological

Antipsychotics

Kim et al., (2014) trialled quetiapine XR (300mg) in a sample of patients recruited from inpatient and outpatient clinics who were experiencing major depressive episodes. Quetiapine XR was administered using an open-trial design against lithium in the treatment of bipolar depression and sleep.
Sleep outcomes assessed using Actigraphy indicated that with quetiapine treatment SE was significantly improved towards the end of the treatment phase (at weeks 6 and 8). WASO was significantly decreased compared with baseline but only at the end of treatment. SOL did not improve significantly over the trial period. There were no significant changes on any objectively measured sleep parameter for patients in the lithium group.

The lithium group showed no significant changes in PSQI score from baseline. Outcomes from the quetiapine group on the other hand demonstrated significant reductions in PSQI scores at weeks, 1, 2, 4, 6 and 8 compared with baseline and the difference between baseline and endpoint yielded an effect that was small-moderate (see Table 2). Although, while a statistically significant, between-group difference was observed at the end of treatment this study may not be sufficiently powered to detect a true effect given a sample of \( n=29 \). In addition, it should be noted that the quetiapine group demonstrated significantly greater sleep impairment than the lithium group at baseline on the PSQI (M 11.4, SD 2.9 vs M 8.4, SD 3.8).

**Ramelteon**

The melatonin-receptor agonist, ramelteon, was trialled to improve sleep for patients with Bipolar Disorder. Studies by McElroy et al., (2010) and Norris et al., (2013) had similar, high scores for methodological quality (scores of 25 and 24 respectively).

McElroy et al., (2010) recruited patients to the trial who had clinically significant sleep disturbance, with a score of >5 on the PSQI. PSQI scores at baseline were not significantly different between the groups (M 14.3, SD 0.79 for the ramelteon group and M 12.9, SD 1.0 for the placebo group). Outcome measures consisted of the PSQI and Pittsburgh Insomnia Rating Scale (PIRS; Moul et al., 2002). The results indicated that patients receiving adjunctive
ramelteon had a similar rate of reduction in insomnia severity to patients receiving the placebo. There were no significant differences between the two groups at end-point although it should be noted that with a sample size of $n=21$, this study may not have been sufficiently powered to detect a significant difference. The observed effect size between baseline and endpoint was small-medium and suggests that ramelteon was actually, less efficacious than placebo. Overall, ramelteon was not associated with significant improvement for any sleep-related outcome variable and was not found to be superior to placebo in reducing symptoms of insomnia after an 8-week intervention period.

Similar results were observed by Norris et al., (2013) who found only a marginal difference on PSQI score between ramelteon and placebo groups after 24 weeks of intervention. Norris et al., recruited a larger sample of $n=83$ patients with existing sleep disturbance using a score of $>5$ on the PSQI as criterion for inclusion, yet, also excluded participants who had diagnoses of primary insomnia disorders. This is unusual given the PSQI is sensitive and specific for insomnia disorder (Buysee et al., 1989). Mean baseline scores for the PSQI were 10.88 (SD 3.44) for the placebo group and 10.17 (3.39) for the ramelteon group. The consistent findings from two studies trialling adjunctive ramelteon would suggest that this drug may not be effective at improving subjective sleep in bipolar disorder.

3.2.3 Effective versus Ineffective Interventions for the Treatment of Sleep

For the purpose of this section studies have been grouped into successful or unsuccessful interventions with successful studies broadly recognised as those that demonstrated improvement in sleep for patients with SMI.

Seven studies (Freeman et al., 2015; Kluge et al., 2014; Luthringer et al., 2007; Yamashita et al., 2004; Kumar et al., 2007, Shamir et al., 2000 and
Tek et al., 2014) could be recognised as effective, illustrating improvements to sleep variables following sleep treatment in patients with schizophrenia. These studies varied in methodological quality ranging from 27 (highest – Tek et al., 2014) to 13 (lowest – Shamir et al., 2000). Tek et al., (2014) demonstrated clear reporting, thorough review and controls for internal and external validity, whereas less methodological scrutiny was applied to these domains in the study by Shamir and colleagues (2000). Of these 7 studies, five specifically recruited individuals with schizophrenia who either complained of sleep problems (Kumar et al., 2007) or formally met criteria for insomnia disorder (Freeman et al., 2015; Luthringer et al., 2007; Shamir et al., 2000 and Tek et al., 2014) providing evidence for the effectiveness of these interventions for patients with existing sleep disturbance. Two studies also noted that their greatest observed improvements were amongst those subgroups with more severe sleep disturbance. Shamir et al., (2000) found that melatonin improved SE to a greater degree for those with below median SE during the placebo period. Yamashita et al., (2004) also found that poor sleep at baseline predicted improvement of subjective sleep quality. Taken together these findings may suggest that the degree of sleep impairment may be related to degree of improvement following sleep intervention.

Wamsley et al., (2013) excluded patients with sleep disorders and did not find any treatment effect following 1 night of eszopiclone treatment. While the brief duration of intervention may provide some explanation for this (particularly given Tek et al., 2014 found improvement to sleep following longer eszopiclone treatment), absence of treatment effect may also be a consequence of low severity of sleep disturbance in the Wamsley et al., (2013) sample. Therefore, some interventions may have limited effectiveness where symptom severity is low, although further research is needed to understand this further.
Goder et al., (2008) similarly did not find a treatment effect for olanzapine. Although it should be noted that both Goder et al., (2008) and Wamsley et al., (2013) were not intended as therapeutic trials. Both studies demonstrated moderate levels of methodological quality with scores of 18 (Goder et al., 2008) and 22 (Wamsley et al., 2013). While Goder et al., did observe increase in SWS following 1 night of olanzapine treatment, they did not observe improvements to subjectively reported sleep, although this may be due to issues with the measure used. The PSQI should be rated over a period of 1 month and thus is more suitable for pre- and post-assessments when the intervention period is longer than the 1 night of intervention described by Goder and colleagues. Yamashita et al., (2004) and Kluge et al., (2014), in their studies that were, respectively, of equal and higher methodological quality, did report improvements to sleep following olanzapine treatment. Furthermore, these studies demonstrated improvements to sleep that were reported using the PSQI following 6 (Kluge et al., 2014) and 8 (Yamashita et al., 2004) weeks of olanzapine treatment respectively. Therefore, it may be that a longer period of treatment is perhaps necessary in order to record outcomes at the subjective level i.e. while olanzapine might quite quickly, following administration, contribute to physiological changes that can be recorded objectively (increase in SWS, observed by Goder et al., 2008), the drug may need to be administered for a longer duration in order for patients to recognise and report changes subjectively. This might explain why Goder et al., (2008) detected changes in sleep across PSG parameters but not on the PSQI.

In studies reporting on sleep treatments in the context of Bipolar Disorder, two studies were successful in demonstrating improvements to sleep; Harvey et al., (2015b) trialling CBTI-BP and Kim et al., (2015) trialling quetiapine XR. The other two studies that failed to identify improvements to sleep following adjunctive ramelteon treatment (McElroy et al., 2010; Norris
et al., 2013). However, both McElroy et al., (2010) and Norris et al., (2013) studies were of high methodological quality and reported similar findings, providing some support for the conclusion that adjunctive ramelteon may be ineffective at producing significant improvement to sleep over placebo. Moreover, all 4 studies examining sleep treatments in bipolar disorder reported high levels of sleep disturbance at baseline with McElroy et al., (2010) and Norris et al., (2013) specifying poor sleep quality (score >5 on PSQI) as part of their inclusion criteria. This suggests that even with clinical levels of sleep disturbance, ramelteon failed to improve sleep for patients in the studies by McElroy and colleagues and Norris and colleagues.

4. Discussion

This paper sought to systematically review the evidence for the treatment of sleep disturbance in people with severe mental illness. This novel review discusses both the level and quality of evidence for the treatment of sleep in schizophrenia and bipolar disorder and uniquely integrates evidence from pharmacological and non-pharmacological interventions in this synthesis. Fifteen studies met criteria for inclusion in the review. These comprised eleven studies of sleep treatment in schizophrenia and 4 studies of sleep treatment in bipolar disorder. All studies were controlled trials reporting on an intervention targeted towards improving sleep parameters and these included a combination of objective and subjective sleep-focused outcomes.

Studies were included if they sampled patients with severe mental illness, although specific sleep complaints were not specified so that the review could broadly include a range of sleep complaints common in SMI. However, this resulted in the inclusion of some studies that did not clearly report the nature or degree of sleep disturbance within their sample. Where
studies did report on the nature of sleep disturbance these largely focused on treatments for insomnia disorder specifically. Other sleep disorders that are prevalent in SMI populations, such as circadian-rhythm sleep-wake disorder and hypersomnia were not well represented in the literature. This is unusual given the high levels of co-morbidity between sleep disturbance and SMI. Of the 8 studies that did provide a summary of their sample’s sleep characteristics, only Harvey et al., (2015b) reported on hypersomnia and circadian-rhythm sleep-wake disturbances in their sample of patients with bipolar disorder. It may be that there is more focus on the treatment of insomnia because this is recognised as the most prevalent sleep complaint in SMI (Cohrs, 2008; Freeman, et al., 2009; Palmese et al., 2011; Xiang et al., 2009), although this creates a bias in the literature over other sleep difficulties that are also prevalent in SMI. That other sleep disorders are under-represented is not to say that they are not present; this may instead be a reflection of under-reporting. As such future studies should seek to address a wider range of sleep disturbance in SMI. Robust and clear approaches to screening patients into trials might also help to improve the extent to which findings can be generalised to similar groups.

The studies that were identified for review varied in sample size and methodological quality. Two studies with the smallest sample size and lowest quality ratings (scores of 11 and 12 for Dursun et al., 1999 and Leibenluft et al., 1997, respectively) were excluded and these are discussed further below. Of the studies that were included however, the smallest sample was n=19 (Shamir et al., 2000) which incidentally also had the lowest quality rating. Even those studies with larger samples e.g. Freeman et al., (2015) with n=50 were not always sufficiently powered. Given the variance in methodological quality it is important that the findings of each study should be evaluated in the context of these parameters.
Non-pharmacological intervention studies (Freeman et al., 2015; Harvey et al., 2015b) were high in methodological quality and both demonstrated improvements in subjective sleep quality, which were maintained at follow-up. Both studies however had small samples and were not sufficiently powered to detect anything but the largest effect size. Also, it was clear from Freeman et al., (2015) and Harvey et al., (2015b) that reductions in insomnia severity occurred across both experimental and comparator conditions. A limitation of these studies is the absence of rigorous control groups against which relative improvements could be considered. Despite these limitations, however, both non-pharmacological intervention trials did demonstrate improvements to sleep, which were also sustained over time. That CBTI improves sleep in schizophrenia and bipolar disorder is also consistent with existing research findings for other mental health conditions (e.g. Harvey et al., 2014; Morin et al., 1994; Myers et al., 2011; Trauer et al., 2015).

Both non-pharmacological interventions included in the review trialled CBT, albeit with some variance in the structure of therapeutic modules, as appropriate to the patient samples studied. In light of the findings from these studies it could be theorised that therapeutic mechanisms for change may occur at the behavioural, affective and cognitive levels. However, given that the two conditions present differently, in that schizophrenia is a psychotic disorder and bipolar disorder is more commonly recognised as an affective disorder (although sometimes with psychotic symptoms), it may be that the mechanisms responsible for changes in sleep, and indeed symptoms, may vary between the two conditions. Further research is needed to explore these mechanisms in detail; for example, future studies evaluating sleep treatment in schizophrenia may not only help to elucidate potential cognitive mechanisms underlying sleep but
may also facilitate understanding of common cognitive mechanisms underlying both sleep and psychosis.

The outcomes from the pharmacological interventions for sleep treatment in schizophrenia suggest that some atypical antipsychotic drugs may be efficacious in treating sleep. Although these findings are often based on individual studies of different medications as it was not often the case that there were multiple studies testing the same drug, with the exception of trials for olanzapine (e.g. Goder et al., 2008; Kluge et al., 2014; Yamashita et al., 2004) and melatonin (Kumar et al., 2007; Shamir et al., 2000) in schizophrenia, and trials of ramelteon (McElroy et al., 2010 and Norris et al., 2013) in bipolar disorder. Even where multiple trials exist there was much variance in study quality or sample, against which it is difficult to draw firm conclusions about the efficacy of particular interventions.

The outcomes from pharmacological interventions for sleep treatment in schizophrenia were however, generally positive, indicating that some atypical antipsychotic drugs, melatonin and hypnotics may contribute to improved subjective sleep quality for patients with schizophrenia. Although the set of studies that sampled patients with bipolar disorder was smaller, there was some evidence to suggest that antipsychotic quetiapine may lead to sleep improvements (notably with small-moderate effect sizes), while adjunctive ramelteon does not.

Unfortunately, no pharmacological trial with either schizophrenia or bipolar patients provided follow-up data which makes it difficult to determine the effectiveness of pharmacological interventions on sleep over time and whether any long-term effects are comparable to those reported for non-pharmacological interventions at follow-up, which is a clear limitation.

Even when pharmacological studies did not find between group differences they did generally note a trend toward improvement to subjective sleep variables for all study participants, irrespective of condition.
It is likely that being part of a sleep study and engaging in assessments and monitoring activities (e.g. sleep diaries) may contribute to placebo effects, which are commonly observed in insomnia trials (Perlis, McCall, Jungquist, Pigeon & Matteson, 2005). For example, Tek et al., observed improved sleep, according to diary ratings, in both placebo and eszopiclone conditions. Shamir et al., (2000) also observed that SE was markedly improved for patients whose SE was low during the placebo arm of the trial, which might be an example of regression to the mean. While such effects are difficult to control in insomnia studies, there might be benefit to sleep as a result of monitoring alone.

Sleep outcomes were reported both objectively and subjectively across the papers included in this review. Although this variance posed a challenge for being able to compare outcomes it was interesting that there was some consistency across studies as to which measures were sensitive to change. For example, Tek et al., (2014) and Freeman et al., (2015) both observed significantly reduced ISI scores at baseline when compared to their respective comparative groups. Both studies also used a sleep diary and found that sleep diary data relating to TST, WASO and SOL also both improved without reaching clinical significance. Sleep diaries can be a useful measure of sleep and are often regarded as the gold standard for subjective sleep assessment (Miller, Kyle, Melehan & Bartlett, 2014), particularly in insomnia (Carney et al., 2012). However, it may be that sleep diaries are perhaps less sensitive than global measures at detecting change. Alternatively, it could be argued that global measures such as the ISI and PSQI demonstrate improvement because they also assess functioning and sleep-related distress, where sleep-diary items do not. This highlights the scope of sleep interventions as perhaps being responsible for improvement across domains related to functioning.
In summary, the findings suggest that both pharmacological and non-pharmacological interventions can be used to treat sleep in schizophrenia and bipolar disorder, although this is based on a small number of intervention trials which varied in methodological quality. A general trend towards improvement was observed across experimental and control groups however and therefore improved procedures that control for placebo effects are required in future studies.

Robustness of Narrative Synthesis and Limitations

Two studies that were identified in the search were excluded from the synthesis: Dursun et al., (1999) and Leibenluft et al., (1997). For transparency, the data extracted from these studies is presented in Table 1 and the quality ratings presented alongside those of other studies in Appendix B. The evidence reported in Dursun et al., was not deemed reliable and thus any conclusions drawn from it may have introduced bias in this review. The study by Leibenluft et al., could also not be considered within the synthesis for reasons pertaining to lack of clarity around the randomised sample; specifically, unclear monitoring of patient attrition and associated incomplete data reporting for an already very small sample of 5 females with rapid-cycling bipolar disorder. Although not fully transparent, it seems that one or more patients had unexpected and adverse reactions to melatonin (and its withdrawal) and were withdrawn from the study, with associated data not available for analysis. It might have been useful for these data to have been presented as a series of case studies instead so that more comprehensive data could be reported for this small group. Given the potential range and severity of reactions, it may also be the case that the findings represent a potentially heterogeneous sample. This was the only study that sampled a group of patients who were rapid-cycling and therefore comparisons cannot be drawn between other studies that recruited similar
samples. The ambitious approach to sampling rapid-cycling bipolar patients provides a clear example of the potential challenges in trialling interventions in complex clinical disorders. Consequently, there perhaps needs to be impetus to improving safety in future clinical trials.

The narrative synthesis consistently extracted and reported information across studies in a similar fashion, highlighting similarities where these occurred. Although every effort was made to ensure a robust synthesis, a number of limitations are acknowledged. Firstly, it should be acknowledged that data extracted from primary studies was completed by one person (author GP), although an independent rater did provide quality assessment ratings to reduce bias in this domain. Another challenge encountered in this review was the synthesis of a variety of outcome data, differing patient samples (where sample descriptions were provided), and range of interventions with varying treatment durations and dosages, for example. It is therefore difficult to draw firm conclusions given the variety of samples, methods and interventions reviewed.

4.1 Clinical Implications
The review has highlighted that sleep can be improved for patients with schizophrenia and bipolar disorder, who present with sleeping difficulties. Patients with SMI should be offered treatments that aim to treat sleep disturbance and these may include pharmacological or non-pharmacological interventions, depending on individual need and resource availability. While non-pharmacological interventions are more labour-intensive, they generally have fewer adverse effects, compared with pharmacological interventions. The non-pharmacological interventions reviewed in this study also produced significant and large improvements to sleep, which were maintained at follow-up. There was some indication that patients with more severe sleep disturbance demonstrate greater
improvement following treatment and this may help prioritise individuals where resources are limited.

4.2 Future Directions

More studies that are rigorously controlled and adequately powered are now needed to develop the findings presented in this review. This review has highlighted a relative dearth of studies examining other common sleep complaints in relation to hypersomnia, circadian-rhythm sleep-wake disorder and nightmare disorders and so trials examining the efficacy of intervention for these conditions should be considered a priority. Future studies should also aim to use clinically relevant outcomes measures (e.g. those recommended by Buysse, Ancoli-Israel, Edinger, Lichsten & Morin, 2006) and triangulate objective and subjective assessments to improve the reliability of reported outcomes. In addition, it would be useful to more fully explore whether severity of sleep disturbance at baseline is related to outcome and future studies should consider recruiting patients with varying severity of sleep disturbances, so that the effectiveness of interventions can be considered in this context.

That sleep can be treated in patients with schizophrenia and bipolar disorder may have implications for other aspects of functioning. For example, as previously noted many studies have observed an association between sleep impairment as contributing to distress, fatigue (Waite et al., 2015), psychotic symptoms (Reeve, Sheaves & Freeman, 2015; Mulligan et al., 2016) and even indicative of relapse in bipolar disorder (Jackson, Cavanagh & Scott, 2003). Future research might then be directed to explore whether, and to what extent, sleep improvements confer benefit to other domains of mental health and functioning. Such research pursuits may further improve understanding about causal relationships between sleep and severe mental illness.
5. Conclusions

This review has identified a number of sleep interventions that can be used to successfully treat sleep in patients with severe mental illness. Of note, CBTI interventions were successful in improving sleep in schizophrenia and bipolar disorder, with effects maintained at follow-up. Findings from pharmacological interventions were more varied, which are likely to have been observed given variation across studies in terms of duration of intervention, dosage and methods of assessment (objective vs. subjective changes). Although, it was consistently found that adjunctive ramelteon did not appear to be effective at treating sleep in bipolar disorder while melatonin was effective at improving sleep for patients with schizophrenia. Atypical antipsychotic medications were generally associated with positive increases in sleep with the exception of perospirone. There was an absence of follow-up data with pharmacological interventions which unfortunately limits understanding about the long-term efficacy of these interventions. It should be noted that the studies varied in size and quality and this may limit the generalisability of the findings.

This review has demonstrated that a range of treatment options exist for the treatment of sleep disturbances in schizophrenia and bipolar disorder, with many contributing to improvements to sleep. Although it should be recognised that these generally reflect improvements to insomnia, given the under-representativeness of other sleep disorders in this review. More trials are needed that are adequately powered, well-controlled and that use standardised assessments so that changes in sleep can be assessed on clinically validated measures. The review indicates that there is value in treating sleep disturbance in severe mental illness, although the evidence base is presently small and in need of development.
6. References

*Included in systematic review


Investigating the effects of poor sleep on cognitive mechanisms underlying psychotic experiences: A comparison of good and poor sleepers.

Prepared for submission to Journal of Abnormal Psychology

Keywords: Psychosis, sleep, insomnia, hallucinations, delusions

Abstract

It is widely recognised that there is an association between sleep and severe mental illness. Emerging evidence suggests that sleep disruption may be causally related to both the occurrence and severity of psychotic symptoms. However, little is known about how this occurs. This study aims to investigate the effects of poor sleep on specific cognitive processes associated with psychotic symptoms, namely hallucinations and delusions.

A non-clinical sample comprising good sleepers (n=39) and poor sleepers (n=22) from the general population were compared on a range of tasks that measure proneness to hallucinations and delusions. In addition, self-reported sleep variables obtained from a sleep diary (retrospectively rated) for the night prior to testing were correlated with computer tasks to examine the relationship between subjective sleep experience and task performance.

The results indicated that there were no significant differences between good sleepers and poor sleepers on any measure, nor was there a significant association between sleep experience prior to testing and task performance. The findings are in contrast to what was expected, although the study was underpowered. A number of methodological issues are also discussed relating to sample characteristics and tests used in the study, and the findings are discussed in this context.
1. Introduction

Sleep problems are a common complaint in people with mental health problems. Prevalence rates for sleep disturbances are as high as 90% in those with depression (Tsuno, Besset & Ritchie, 2005). Insomnia, which is characterised by difficulty initiating and maintaining sleep, is the most commonly reported form of sleep disturbance across psychotic disorders. In addition, approximately 40% of all insomnia patients are reported to experience a co-existing mental health problem (McCall, 2001) which indicates a high degree of comorbidity.

Sleep disturbance is recognised as a core feature of Bipolar Disorder (BD) (Harvey, Schmidt, Scarna, Semler & Goodwin, 2005). A recent review by Kanady, Soehner & Harvey (2015) observed high levels of co-occurring sleep disturbance in BD: manic periods often consisted of reduced need for sleep and insomnia; and depressive periods often consisted of hypersomnia and insomnia.

Alfonso, Brissos, Figueira & Paiva (2011) also found a trend towards more disrupted sleep-wake cycles for those individuals with schizophrenia who had mainly positive symptoms compared to those with mainly negative symptoms. Moreover, sleep disruption has been associated with increased severity of psychotic symptoms (Mulligan, Haddock, Emsley, Neil & Kyle, 2016; Waters et al., 2011) and specifically, positive symptoms of psychosis (Davies, Haddock, Yung, Mulligan & Kyle 2016).

There is consistent evidence for impaired sleep as contributing to the severity of psychotic symptoms and within this there is some indication of causality, although this has yet to be empirically tested. Although, it seems that both degree of sleep disturbance and severity of psychiatric symptomatology exist on a continuum. Indeed, similar observations between sleep disturbance and psychotic symptoms have been evidenced in non-
clinical samples. For example, data from a comorbidity sleep survey replication (Roth et al., 2006) showed that adults from the general population who reported specific sleep problems e.g. difficulty falling asleep/waking up, were at greater odds of reporting psychotic experiences over the past year, even after adjusting for DSM-IV comorbid problems.

Petrovsky et al., (2014) observed that sleep deprivation in healthy adults mimicked neurological processes underlying psychosis. Petrovsky et al., found that pre-pulse inhibitions, which prevent interruption of perceptual and sensory analysis, which have been shown to be consistently reduced in schizophrenia, were also reduced in a sample of healthy adults who had undergone a night of sleep deprivation. Petrovsky and colleagues found that sleep deprivation actually induced some psychotic-like experiences such as perceptual distortion, cognitive disorganisation and anhedonia, although not delusions, mania or paranoia. The authors argued that perhaps the lack of association with delusions and paranoia did not emerge during the study because of the relatively small duration of sleep deprivation. In another sleep-deprivation study Kahn-Greene, Killgore, Kamimori, Balkin & Killgore (2007) found that two-nights of sleep deprivation contributed to an increase in self-reported affective symptoms (e.g. anxiety, depression) and changes at the neurobiological level, suggesting that sleep deprivation may disrupt affective processing.

Research evidence also indicates that in addition to affective processing, sleep disturbance can also affect cognitive processing. Findings from a meta-analysis indicated that insomnia was related to impaired performance on tasks related to episodic memory, problem-solving, manipulation and retention in working memory, but not attention, perceptual and psychomotor processes and general cognitive functioning (Fortier-Brochu, Beauliu-Bonneau, Ivers & Morin, 2012). This suggests that insomnia may affect specific, as opposed to global, cognitive processes.
Fortier-Brochu, Beaulieu-Bonneau, Ivers & Morin (2012) concluded that despite insomnia contributing mild to moderate impairments on memory function, these findings were nevertheless consistent with other studies evidencing impaired memory function in insomnia (Drummond, Walker, Almklov, Campos, Anderson & Straus, 2013; Lim & Dinges, 2010; Yoo, Hu, Gujar, Jolesz & Walker, 2007) including those based on subjective self-report (Roth & Ancoli-Israel, 1991).

These findings consistently demonstrate that loss of sleep can contribute to disrupted processing across both affective and cognitive domains and furthermore, that such disruption is associated with symptoms of psychosis. However, there has been little examination of how sleep affects the specific cognitive processes that underlie symptoms of psychosis. In particular, whether and to what extent sleep may affect cognitive biases which underlie hallucinations and delusions.

Delusions are strongly held beliefs, often with referential, persecutory or grandiose content and are key characteristics of serious mental illnesses like schizophrenia and bipolar disorder, although they have also been reported in non-clinical samples (Johns et al., 2004; Kahn-Greene et al., 2007; Wiles et al., 2006). Delusions represent a data-gathering bias in that they are thought to occur as a result of making (biased or faulty) judgements based on limited data. Reduced data-gathering, also known as ‘jumping to conclusions’ (JTC) is a pattern of thinking that has been repeatedly identified in samples experiencing delusions, but has also been associated with working memory impairment (Freeman et al., 2014).

Jumping to conclusions is often assessed using a probabilistic reasoning task, often referred to as the beads task (e.g. Garety et al., 2005; Menon, Mizrahi & Kapur, 2008; Peters, Thornton, Sikou, Linney & MacCabe, 2008) in which individuals are required to make decisions on the basis of limited information. The key variable in this task is the number of
stimuli needed to make a decision and Garety et al., (2005) suggests that requesting two or fewer beads is indicative of the JTC bias. Dudley, Taylor, Wickham & Hutton (2015) completed a meta-analytic review of the JTC bias and found that the JTC bias was associated with greater probability of delusional occurrence.

Hallucinations have been hypothesised to represent an externalising bias i.e. a tendency to misattribute internally generated experiences to external sources. This externalising bias has been demonstrated using auditory signal detection tasks (SDT) in which individuals are required to respond to voices they may have heard under ambiguous sensory conditions i.e. voices heard amidst a background of white noise. Auditory signal detection tasks have consistently been used to demonstrate the presence of this externalising bias amongst clinical (e.g. Varese et al., 2012) and non-clinical samples (e.g. Barkus, Stirling, Hopkins, McKie & Lewis, 2007; Barkus et al., 2011) illustrating that it is a robust indicator of this phenomenon (Brookwell, Bentall & Varese, 2013).

Barkus, Stirling, Hopkins, McKie & Lewis (2007) used a SDT that required participants to respond ‘yes’ or ‘no’ to indicate whether they had perceived hearing a voice played amidst a background of white noise. Barkus et al., recorded responses (hits, false alarms, correct rejections and misses) as well as reaction times (speed of responding) and found that individuals with high hallucination proneness not only reported more false alarms (making a ‘yes’ response in the absence of a voice being played) than those with moderate and lower levels of hallucination proneness, but also that high hallucination-prone individuals also made their responses more quickly. This rapid decision making about ambiguous stimuli may provide evidence to suggest that the JTC bias commonly found amongst people prone to delusions may also be present in those prone to hallucinations. Moreover, given that hallucinations and delusions often co-vary, an
assumption could be made that common mechanisms underlie both hallucinatory and delusional experience, yet this remains little explored to date.

That sleep disturbance has been associated with hallucinations and delusions in clinical and non-clinical groups highlights two main issues. Firstly, that there are cognitive and affective mechanisms underlying psychotic experience that are common to both clinical and non-clinical groups. Secondly, that sleep may influence psychotic-like experiences through impacting on these key cognitive and affective mechanisms. However, no study has yet assessed whether sleep disturbance is linked to alterations in cognitive process, particularly those which are presumed to confer risk for psychotic like experiences, specifically cognitive biases.

1.2 Aims

This study aimed to investigate the relationship between sleep and cognitive biases that have been associated with hallucinations and delusions. This was achieved by comparing good and poor sleepers (using a model of insomnia) in the absence of psychiatric co-morbidity, on validated assessments and probes of psychotic-like experiences.

The primary aim of this study is to investigate how good and poor sleepers differ in performance on cognitive tasks that have been associated with psychotic symptoms.

*Hypothesis 1*: There will be a significant difference between good sleepers and poor sleepers across performance on all computer tasks and questionnaires measuring proneness to hallucinations and delusions. Specifically:

a) Poor sleepers will report greater proneness to hallucinations compared to good sleepers
b) Poor sleepers will report more unusual experiences than good sleepers

c) Poor sleepers will report more paranoid thoughts than good sleepers

d) Poor sleepers will demonstrate poorer working memory than good sleepers

e) Poor sleepers will report more false alarms than good sleepers on a voice-hearing task

f) There will be more poor sleepers evidencing the JTC bias than good sleepers.

The secondary aim of this study is to assess whether sleep prior to testing is associated with task performance the following day.

*Hypotheses 2*: Subjectively reported sleep experience for the night prior to testing will be associated with task performance:

a) Hours of sleep, sleep efficiency and sleep quality will be negatively correlated with frequency of false alarms on the voices task.

b) Hours of sleep, sleep efficiency and sleep quality will be positively correlated with working memory performance.

### 2. Method

The study received ethical approval from the University of Manchester Research Ethics Committee (Ref: 15290).

#### 2.1 Design

The current study used a between-subjects design. Two groups (good sleepers vs. poor sleepers) were compared on a series of cognitive tasks. Participants’ sleep, subjectively rated from the night before testing, was also investigated as a predictor of task performance.
2.2 Participants

An opportunistic sampling methodology was used to recruit volunteers; recruitment posters were placed in buildings in and around the University of Manchester, UK and Manchester City Centre and advertisements for participants were also circulated on internal research bulletins and across social media.

2.3 Procedure

2.3.1 Online Screening

Participants expressing an interest in the study were directed to complete online screening questions to determine their eligibility for the study. The screening webpage was open from October 2015 until the end of April 2016.

Screening questions included self-report assessment in relation to the study’s inclusion criteria, followed by completion of up to 4 further questionnaires.

Inclusion Criteria

The minimum inclusion requirements stipulated that participants:

1. were aged between 18-65 years
2. met criteria for inclusion within Poor Sleep Group OR Good Sleep Group (see below).
3. self-reported that they were not currently receiving, or had recently received, treatment for a mental health problem
4. self-reported that they were not taking prescribed sleeping medication
5. self-reported they were not using any psychoactive substances (substances that affect mood, perception, behaviour e.g. cannabis, other illegal drugs or legal highs) and had not been in treatment for drug or alcohol related difficulties in the past 12 months.
Criteria 3-5 were specified to ensure that a clearly defined non-clinical sample was recruited to the study. Participants were also excluded if they worked shifts or if they had a neurological/physical impairment that prevented them from engaging with the computerised tasks.

**Screening Questionnaires**

Participants who met the inclusion criteria for the study were directed to complete up to 4 further screening questionnaires. Each questionnaire was scored immediately, as soon as data were submitted, which enabled a prompt assessment of whether the exclusion threshold had been met. When the exclusion threshold was met, participants were notified through presentation of an automated message to inform they had been screened-out of the study. Participants who were screened out were not required to complete any further questionnaires. Only participants who continued to score within the specified inclusion limits were permitted to progress through the screening questionnaires.

**Depression Anxiety and Stress Scale (DASS; Lovibond & Lovibond, 1995)**

The DASS is a widely used tool for the assessment of depression, anxiety or stress. The 21-item version of the DASS was used in this study and participants scoring within the severe range for either depression (≥20) or anxiety (≥14) were excluded.

**Sleep Disorders Algorithm (Wilson et al., 2010)**

The sleep disorders algorithm developed by Wilson and colleagues (2010) is a brief questionnaire that can be used to screen for the presence of

---

3The original cut-off score for exclusion was a score within or above the moderate range on either depression or anxiety scores (score ≥10 anxiety; ≥14 depression). However, due to high rates of exclusion this criterion was revised to exclude persons scoring within or above the severe range on either subscale.
any other sleep disorder that isn’t insomnia. Five lead questions prompt assessment of the following sleep disorders: narcolepsy, sleep breathing disorder, periodic leg movement syndrome/ restless leg syndrome, circadian rhythm sleep disorder and parasomnia. Supplementary questions are only asked when the response to the lead question is positive, to determine potential presence of sleep disorder. Individuals identified as having sleep disorders other than insomnia were excluded.

Pittsburgh Sleep Quality Index (PSQI; Buysse, Reynolds, Monk, Berman & Kupfer, 1989).

The PSQI is a widely used instrument for the assessment of sleep quality and disturbance. The measure loads onto 7 dimensions: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication and daytime dysfunction. The 7 component scores contribute to a global PSQI score ranging from 0 to 21 with a higher score indicating greater impairment. Buysse and colleagues (1989) suggest a score of >5 to distinguish poor sleepers. Participants achieving a PSQI score of ≥6 were considered eligible for the poor sleep group in this study to ensure a robust poor-sleep group. Participants scoring ≤3 were considered eligible for the good sleep group, providing these individuals also had a minimum of 7 hours of sleep per night (extracted from their response to question 4 of the PSQI).

Sleep Condition Indicator (SCI; Espie et al., 2014)

The SCI is a brief, 8-item questionnaire used to evaluate insomnia disorder. The assessment has good convergent validity with the PSQI and provides a useful analysis of insomnia symptoms against criteria detailed in the DSM-V (American Psychiatric Association, 2013). Scores range from 0-32 with a higher score indicating better sleep. Only participants scoring ≥6 on
the PSQI (poor sleepers) were required to complete the SCI to confirm their status as a poor sleeper for inclusion within that group. Therefore, those individuals who scored ≥6 on the PSQI and ≤16 on the SCI satisfied inclusion criteria for the poor sleep group.

Main Testing Phase

All participants who successfully progressed through the screening questions, and therefore met eligibility criteria for the study, were invited to attend the University of Manchester to complete the main study tasks. Each participant was tested on an individual basis and testing took place in a private research cubicle, located off a quiet corridor in a University building. All participants provided informed consent to participate prior to engaging with the tasks. During the 40-minute testing session, each participant completed three computer tasks then completed a sleep diary and questionnaire pack.

2.4. Materials

2.4.1 Computer Tasks

The computer tasks were administered first and the order of these 3 tasks was counterbalanced between participants in a random order. The computer tasks were performed on a Dell desktop computer and tasks operated from E-Prime version 1.

N-back

The N-back task provides an assessment of working memory function (Gervins & Cutillo, 1993). The task involves presentation of stimuli and requires on-line monitoring, updating and manipulation of information held

---

4 Individuals who scored >16 on the SCI were initially excluded. However, due to extremely low recruitment rates a decision was made to remove the SCI criterion part way through recruitment and base group allocation solely on the PSQI criterion (score of ≥6).
in working memory. This task can be delivered electronically which has the advantage of minimising bias, unlike other methods of working memory assessment (e.g. digit span assessments).

As part of the n-back task participants were presented with a series of letters that each appeared individually on the screen in front of them. The pseudorandom sequence of capitalised letters were presented centrally for a duration of 500ms and 2500ms stimulus interval (taken from Braver et al., 1997).

After the presentation of the third letter in the sequence, participants were asked to make a response as to whether the letter presented (target) matched the letter that appeared two letters earlier in the sequence, consistent with the 2-back variant of this task. Fifty letters were presented in each block, rendering a total of 48 responses. When the target letter matched the letter from 2-before, participants were instructed to press M on the keyboard to record a match and to press Z to record each no-match. Matches were set to occur at a rate of 30%.

The task was presented in 6 blocks, each lasting 2 minutes and 30 seconds. The first two blocks functioned as practice blocks to orient participants to the task. Participants were provided with feedback in the form of percentage correct, after each practice block. Four main-task blocks followed without feedback in-between. Scores for this task were expressed as a percentage of correct ‘matches’ per block. Performance across the 4 main-task blocks was averaged to calculate average performance. A higher score on this task was illustrative of better working memory performance.

Auditory voice detection bias task

An auditory voice detection task was used to measure propensity for false perceptions. Participants were provided with good quality headphones (Sennheiser HD 215) and were asked to wear these for the duration of the
voice detection task. The task was developed by Huque, Poliakoff & Brown (2016) and involved responding to voices amidst a background of white noise. Participants were asked to respond to voices they heard by pressing the space-bar on the keyboard.

Voices were created using text to speech software that converted nonsense words (each 7-letters long) into male voice-clips. Voice-clips that were identified as sounding very dissimilar to English words were deliberately selected for use.

The task consisted of a practice phase (approximately 1 minute) followed by main phase (4 minutes 30 seconds). The purpose of the practice task was to a) familiarise participants to the task and b) determine reaction time criteria to identify hits and false alarms in the main task phase.

During the practice phase 24 voice clips were played each for 800ms at random intervals, ranging from 1000ms – 2000ms seconds between each sound file. Eighteen volume levels were used and these amplitudes ranged from -1000 - -6200 (amplitude units as specified by E-Prime). Voices heard during the practice task therefore consisted of loud voices and whispers.

For the main task phase participants were informed that they would only hear whispers and to respond to these as they had done previously. The main task consisted of 18 voice-clips at amplitudes of between (-5450 and -6200) at random intervals ranging between a minimum gap of 3 seconds and maximum gap of 10 seconds between two voices.

The data generated from this task (practice and main phase) consisted of hits and false alarms. As the programme generates psychophysical data, the cause for a response cannot always be accurately recognised as a hit or false alarm and so the data were processed using logical calculations, with response times (as opposed to responses themselves) used to identify hits and false alarms, following guidance from the programme developers (Huque, Poliakoff & Brown, 2016). The task measured response times
(milliseconds, ms) and these data were used to determine the accuracy of perceptions (hits and false alarms). Firstly, mean response-time data were calculated using data obtained from the practice task. Upper and lower-limits for reaction times were then calculated using first and third quartiles respectively. Response times that fell outside of these limits were identified as false alarms. The false alarm rate was taken to illustrate a tendency towards hallucinatory experience. Therefore, increased frequency of false alarms was taken to indicate greater predisposition to externalising bias.

**Beads Task**

The beads task is a probabilistic reasoning task which is thought to be a good indicator of a data-gathering bias that is often associated with delusions (Garety & Freeman, 2013; Fine, Gardner, Craigie, & Gold, 2007; Freeman, 2007). Garety et al., (2013) suggests that the selection of two or fewer beads to reach a decision is indicative of the JTC bias and this has been observed to be prevalent in clinical samples at 27% (Garety et al., 2013) and non-clinical samples at 20% (Freeman, Pugh & Garety, 2008). These rates have been established using the more difficult version of the task (bead ratio of 60:40) which has also been selected for use in this study.

A computerised version of the beads task was used in this study. Participants were presented with a display of two jars, each containing 100 beads. Each jar contained a mixture of red and blue beads in 60:40 ratios; one with 60 red and 40 blue and the other with 60 blue and 40 red beads. Participants were informed that one of the jars had been selected at random by the computer. In order to determine which jar had been selected (mainly red or mainly blue) participants were advised that they could request to see beads from the selected jar – these were presented and then replaced to keep the proportions the same. With every key press, the participant could select to see more beads, which then remained on the screen to act as a memory
aid. Participants were encouraged to make a decision about which jar the beads were being drawn from using the smallest number of beads as possible but that they could request up to 40. Participants selecting two or fewer beads were recognised as demonstrating the JTC bias.

2.4.2 Sleep Diary & Questionnaires

Consensus Sleep Diary, Carney et al., (2012)

As cognitive performance may be modulated by quality of sleep on the night prior to testing (Vallieres, Iver, Bastien, Beaulieu-Bonneau & Morin, 2005), a sleep diary was used to obtain subjective sleep assessment information from participants. The consensus sleep diary (morning version) developed by Carney et al., (2012) was used for participants to retrospectively rate their previous night’s sleep (night prior to being tested at the University). Example items from the morning diary included: how long did it take you to fall asleep, what time was your final awakening, and how would you rate the quality of your sleep? Three variables were extracted from the diary: sleep duration, sleep quality and sleep efficiency. Sleep duration was extracted from item 8 (in total, how long did you sleep). Similarly, sleep quality was extracted from item 9 (how would you rate the quality of your sleep), where participants were asked to provide a likert scale rating of 1 – 5 as to the quality of their sleep, with a higher score illustrative of greater satisfaction with sleep. Sleep efficiency was expressed as a percentage to represent potential time spent asleep, given the time spent in bed.

Launay-Slade Hallucination Scale (LSHS; Launay & Slade, 1981)

The Launay-Slade Hallucination Scale is a 12-item measure that assesses hallucination proneness on a unidimensional scale. Items include
clear-cut hallucinatory experiences as well as sub-clinical experiences aimed to assess a continuum of hallucinatory experience. The scale has good levels of test re-test reliability ($r=0.84$, $p<0.01$; Bentall & Slade, 1985) and validity (75% correct classification using cross-validation (LOOCV), $z = 2.79$, $p<0.01$; Serper, Dill, Kot, Chang & Elliot, 2005).

*Oxford Liverpool Inventory of Experiences and Feelings (O-LIFE; Mason, Claridge & Jackson, 1995)*

The O-LIFE assessment comprises 4 subscales (unusual experiences, cognitive disorganisation, introvertive anhedonia and impulsive non-conformity) that measures psychosis-proneness or psychotic characteristics in non-clinical samples. Only the unusual experiences subscale (O-LIFE-UE) was used in this study to assess positive symptoms of psychosis. The subscale has high internal consistency ($\alpha=0.89$; Mason et al., 1995) and high test-retest reliability ($r=0.86$, $p<0.001$; Burch, Steel & Hemsley, 1998) in non-clinical population samples.

*Green Paranoid Thoughts Scale (GPTS; Green et al., 2008)*

This questionnaire is a multidimensional measure of persecutory ideas developed for use across the general population-psychopathology continuum. The GPTS has two, 16-item subscales which assess ideas of social reference and persecution. The GPTS has high levels of reliability (i.e. internal consistency $\alpha =.90 – \alpha =.95$ for subscale and total scores) in non-clinical samples (Green et al., 2008).

2.5 Data Analysis Plan

An a-priori power calculation indicated that an effect size of 0.4 (based on working memory effect sizes; Fortier-Brochu et al., 2012), alpha set
at 0.05 and assumed power at 80%, the study would require a sample of $n=45$ in each group so this was the target sample size. All statistical analyses were performed on IBM SPSS version 22.

Participant characteristics and dependent variables were assessed for normality via inspection of histograms as well as skewness and kurtosis statistics. All variables were not normally distributed. Attempts were made to correct skewed data using log transformations although this did not significantly improve normality. As such, data analysis proceeded using bootstrapping to correct for the assumption of normality not being met. All confidence intervals reported are therefore bias corrected.

In line with hypothesis 1, the planned statistical analyses included a series of t-tests to assess between-groups comparisons (good sleepers vs. poor sleepers) on all computer tasks (working memory, voice-detection task and beads task) and questionnaires (GPTS, LSHS, O-LIFE-UE).

For hypothesis 2, correlational analyses (using Pearson’s correlation co-efficient) were planned to assess a potential predictive relationship between the previous night’s sleep and performance the following day. This was achieved by extracting the following variables from the sleep diary: hours of sleep, sleep efficiency, and sleep quality, and correlating these against performance on the working memory task and frequency of false alarm responses in the voice-detection task. Sleep diary items were not explored in relation to questionnaire measures as these were thought to represent more stable phenomena, not be amenable to significant variation and thus not thought to be dependent on previous night’s sleep.

3. Results

3.1 Sample Description

Between October 2015 and April 2016, 237 individuals accessed the screening questions of which 79 met criteria for either the good sleep group
or the poor sleep group. Eighteen participants did not attend for further participation (beyond screening) and therefore the final sample size was N=61 (n=39 good sleepers and n=22 poor sleepers). Fifty-three females (86%) took part in the study and the gender distribution across both groups was roughly equal. Descriptive statistics are presented in Table 2.1 below. Mean and median values are reported given data were not normally distributed.
Table 1
Demographic characteristics of good sleepers and poor sleepers

<table>
<thead>
<tr>
<th></th>
<th>Good Sleep Group</th>
<th>Poor Sleep Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n=39 )</td>
<td>( n=22 )</td>
</tr>
<tr>
<td><strong>Gender: (% female)</strong></td>
<td>n=34 (87%)</td>
<td>n=19 (86%)</td>
</tr>
<tr>
<td><strong>Mean (SD)</strong></td>
<td><strong>Mdn</strong></td>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>24.97 (8.57)*</td>
<td>33.45 (15.08)</td>
</tr>
<tr>
<td></td>
<td>22 (9)</td>
<td>29.50 (29)</td>
</tr>
<tr>
<td><strong>Screening Measures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DASS Anxiety subscale</td>
<td>2.35 (2.90)</td>
<td>3.18 (2.87)</td>
</tr>
<tr>
<td></td>
<td>2 (4)</td>
<td>2.00 (6.00)</td>
</tr>
<tr>
<td>DASS Depression subscale</td>
<td>3.08 (3.03)</td>
<td>3.27 (2.72)</td>
</tr>
<tr>
<td></td>
<td>2 (6)</td>
<td>2.00 (2.00)</td>
</tr>
<tr>
<td>PSQI Global score</td>
<td>2.21 (0.98)**</td>
<td>9.00 (2.53)</td>
</tr>
<tr>
<td></td>
<td>3 (1)</td>
<td>9.00 (3.25)</td>
</tr>
<tr>
<td>SCI Total</td>
<td>-</td>
<td>16.36 (6.65)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.00 (10)</td>
</tr>
<tr>
<td><strong>Sleep Diary Items</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep Duration (H:MM)</td>
<td>7:44 (1:03)**</td>
<td>6:49 (1:21)</td>
</tr>
<tr>
<td></td>
<td>7:40 (1:20)</td>
<td>7.00 (1.44)</td>
</tr>
<tr>
<td>Sleep Efficiency (%)</td>
<td>90.10 (7.95)**</td>
<td>81.09 (10.91)</td>
</tr>
<tr>
<td></td>
<td>90.71 (11.94)</td>
<td>83.45 (12.63)</td>
</tr>
<tr>
<td>Sleep Quality</td>
<td>4.26 (0.72)**</td>
<td>3.27 (1.12)</td>
</tr>
<tr>
<td></td>
<td>4 (1)</td>
<td>3 (2)</td>
</tr>
</tbody>
</table>

\*\( p < .05 \), \*\* \( p < .01 \)
There were no significant differences between the groups on gender distribution ($\chi^2(1) = 0.01, p=0.93$) or DASS anxiety ($t(59) = -1.07, p=0.30$) [95% CIs -2.25, 0.75] depression ($t(59) = -0.25, p=0.80$) [95% CIs -1.82, 1.38] subscales. However, participants in the good sleep group were significantly younger ($t(28.82) = -2.43, p=0.02$) [95% CIs -15.70, -2.09].

As would be expected, the groups significantly differed with respect to PSQI score ($t(24.61) = -12.12, p=0.001$), [95% CIs -7.97, -5.77] and sleep diary items. Participants in the poor sleep group had a shorter sleep duration ($t(35.75) = 2.77, p=0.01$), [95% CIs 0:19, 1:34], poorer sleep efficiency ($t(33.78) = 3.39, p=0.01$) [95% CIs 4.05, 14.06] and reduced quality of sleep ($t(30.86) = 3.71, p=0.01$), [95% CIs 0.45, 1.51] compared with participants in the good sleep group.

3.2 Comparing good sleepers and poor sleepers on questionnaires and computer tasks associated with psychotic-like experiences.

Given the difference in age between the two groups age was correlated against all dependent variables of interest (LSHS, O-LIFE-UE, GPTS, working memory and false alarms on the voice-detection task). Pearson’s correlation co-efficient indicated that age was negatively correlated with the ideas of reference subscale from the GPTS ($r=-0.27, p=0.05$). Given the frequency of correlations the alpha was set at $p<0.01$ and accordingly this correlation was not considered significant.

Table 2 illustrates between-group differences for good sleepers and poor sleepers across questionnaire and computer tasks that assess cognitive processes underlying psychotic-experiences.
Table 2
Comparisons between good sleepers and poor sleepers on questionnaires and computer tasks associated with psychotic-like experiences.

<table>
<thead>
<tr>
<th>Questionnaires</th>
<th>Good Sleep Group</th>
<th>Poor Sleep Group</th>
<th>t(53)</th>
<th>p</th>
<th>95% CIs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=39</td>
<td>n=22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD) Mdn (IQR) Mean (SD) Mdn (IQR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O-Life Unusual Experiences Subscale Score</td>
<td>4.27 (3.83)</td>
<td>4 (5)</td>
<td>3.28 (2.56)</td>
<td>2.50 (5)</td>
<td>.99</td>
</tr>
<tr>
<td>Launay-Slade Hallucination Scale</td>
<td>1.89 (1.41)</td>
<td>1 (2)</td>
<td>2.11 (1.68)</td>
<td>2 (3)</td>
<td>-.51</td>
</tr>
<tr>
<td>Green Paranoid Thoughts Scale</td>
<td>40.24 (8.83)</td>
<td>38 (12)</td>
<td>41.17 (15.55)</td>
<td>34.50 (11)</td>
<td>-.28</td>
</tr>
<tr>
<td>(GPTS Total)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPTS Ideas of Reference Subscale</td>
<td>22.68 (6.29)</td>
<td>21 (10)</td>
<td>22.56 (10.19)</td>
<td>18 (10)</td>
<td>0.54</td>
</tr>
<tr>
<td>GPTS Ideas of Persecution Subscale</td>
<td>17.57 (3.80)</td>
<td>16 (2)</td>
<td>18.61 (5.70)</td>
<td>16 (2)</td>
<td>-0.81</td>
</tr>
<tr>
<td>Computer tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-Back performance (%)</td>
<td>87.05 (11.46)</td>
<td>90.11 (7.29)</td>
<td>85.82 (17.01)</td>
<td>91.15 (8.70)</td>
<td>0.32</td>
</tr>
<tr>
<td>(working memory)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of False Alarms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.57 (3.29)</td>
<td>3 (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.94 (3.61)</td>
<td>2 (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.64</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-1.37, 2.21]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. significant if $p<0.05$
As presented in Table 2 there was no significant difference between groups across any questionnaire or computer measure; good sleepers did not differ significantly from poor sleepers on measures that probe psychotic-like experiences. The JTC bias (selection of 2 or fewer beads) was indicated for two participants, both of whom were good sleepers. Given the low frequency of individuals demonstrating the bias overall it was not possible to complete a Chi-square to assess whether the proportion of those demonstrating the bias was significantly different between groups.

3.2 Association between previous night’s sleep and task performance

A series of correlations (using Pearson’s correlation co-efficient) were performed on data for all participants, irrespective of group, to test whether there was an association between previous night’s sleep and working memory and frequency of false alarms (voice-detection task). These data are presented in the correlation matrix presented in Table 3 below. The alpha level was set to $p<0.01$ to adjust for the number of correlations performed.

<table>
<thead>
<tr>
<th></th>
<th>4. N-back performance</th>
<th>5. Frequency of false alarms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sleep Duration</td>
<td>0.20</td>
<td>0.07</td>
</tr>
<tr>
<td>2. Sleep Quality</td>
<td>0.21</td>
<td>0.14</td>
</tr>
<tr>
<td>3. Sleep Efficiency</td>
<td>0.19</td>
<td>0.15</td>
</tr>
</tbody>
</table>

*Note* *significant at $p<0.01$. 

Table 3

Correlation matrix detailing associations between sleep variables and performance on working memory and voice detection task
The results indicate no significant associations between sleep experience prior to testing and working memory performance and false alarms.

4. Discussion

The current study sought to investigate associations between poor sleep and cognitive processes related to the experience of psychotic symptoms. This was achieved by comparing good and poor sleepers on a range of tasks which probe psychotic-like experiences, including cognitive biases that are thought to underlie hallucinations and delusions. The study sampled non-clinical adults to enable a comparison between good and poor sleepers (those meeting criteria for insomnia disorder), in the absence of psychiatric co-morbidity.

The study successfully recruited two well-defined groups of good sleepers and poor sleepers and the significant difference between groups on the sleep diary confirmed that the grouping approach was robust. To contextualise these data, Morin, Beaulieu-Bonneau, LeBlanc & Savard (2005) report mean scores of 9.67 (SD 3.26) for those with insomnia syndrome and Backhaus Junghanns, Broocks, Riemann & Hohagen, (2002) also report a mean PSQI score of 12.5 (SD 3.8) for a sample of individuals with primary insomnia, both studies involving similar non-psychiatric groups. These data would indicate that the poor sleepers recruited to this study had similar levels of sleep impairment to those with insomnia disorder.

The primary hypothesis stated that poor sleepers would report increased proneness to hallucinations and delusions, report more false alarms and have poorer working memory than good sleepers. The hypotheses also anticipated a higher prevalence of the JTC bias amongst the poor sleepers than the good sleepers. However, the results from the current study did not suggest that the groups differed significantly and therefore the null hypotheses could not be rejected.
The secondary hypothesis assumed a relationship between state-sleep experience (on the night prior to testing) and task performance. Correlational analyses performed on data for the whole sample indicated that there were no significant associations between sleep variables, which included hours of sleep, sleep efficiency or sleep quality prior to testing, and performance on the computer tasks that measured working memory (N-back) and attributional biases (false alarms on the voice-detection task).

Results from the current study are in contrast to what was expected, given findings from other studies that have found an association between impaired sleep and psychotic-like experiences (e.g. Jeppesen et al., 2015; Petrovsky et al., 2014; Reeve, Sheaves & Freeman, 2015; Thompson et al., 2015). There are a number of potential explanations for why this study failed to find any significant differences between groups or associations between sleep and psychotic-like experiences. First of all, the study was underpowered and thus may not have been able to detect a significant result if one existed. Secondly, over half of the study sample comprised students and their sleeping patterns may not represent sleep data typical for a general population sample. Thirdly, it may be that tasks selected to measure memory, hallucination and delusion proneness were not sensitive enough. The beads task has been used to identify the presence of the JTC bias in other studies with non-clinical groups (e.g. Freeman, et al., 2008; Startup, Freeman & Garety, 2008) and Freeman, et al., (2008), for example, identified the JTC bias in 20% of their non-clinical sample which is higher than the 3% prevalence established in the current study.

The data from the voices task indicated that there was a low hit-rate overall, which suggests that participants made fewer responses on this task in the current study, compared to other studies (Huque, Poliakoff & Brown, 2016). It was thought that this may have been due to difficulty hearing voices, although further investigation with sound-detection software
indicated that this was not likely, although it should also be noted as a potential limitation.

The negatively skewed data for working memory may also indicate ceiling effects. Although many participants complained about the difficulty of the 2-back version of the task, perhaps a more difficult version of the memory task (e.g. 3-back) might have encouraged more variation in scores. While working memory is often reported to be impaired in sleep deprived groups, the magnitude of this effect sometimes appears to be quite small (e.g. Fortier-Brochu et al., 2012). It could be argued that that sleep may not have been sufficiently impaired in the sample to confer impairment on this task although, as previously reported, the levels of insomnia in this group were comparable to other non-clinical samples, illustrating significant sleep impairment.

This study purposefully used a sample of healthy, non-clinical adults in an attempt to isolate the role of sleep on psychotic-like experiences, in the absence of psychiatric co-morbidity. However, in carefully selecting a non-clinical sample, the study may have inadvertently reduced or eliminated key factors, such as high levels of psychosocial stress, known to be related to psychotic-experiences in clinical samples (Nuechterlin & Dawson, 1984) and also associated with psychotic-like symptoms in non-clinical samples (e.g. Barrantes-Widal, Chun, Myin-Germey & Kwapil, 2013; Lincoln, Peter, Schäfer & Moritz, 2009). The absence of between-group differences may therefore be because the healthy sample of volunteers were perhaps deficient of key characteristics which may well mediate the occurrence of psychotic-like experiences. For example, Mulligan, Haddock, Emsley, Neil & Kyle (2016) recently identified that sleep disturbance predicted hallucinations and delusions but more importantly, that these relationships were mediated by negative affect.
The threshold criteria for anxiety and depression (on the DASS) were increased in this study however, to allow those with greater scores into the study. This means that despite there being some participants experiencing moderate levels of anxiety or depression, there were no differences between groups, indicating that perhaps there were other reasons for non-significant differences between the groups.

Finally, it is also useful to consider the possibility that these findings indicate that there may be no differences between good sleepers and poor sleepers on measures that probe psychotic-like experiences. This, as a conclusion however, is inconsistent with other studies that have demonstrated a link between sleep impairment and cognitive processes (e.g. Fortier-Brochu et al., 2012).

The findings of this study should also be considered in the context of its limitations. Firstly, whilst every effort was made to obtain a sample from the general population, a large proportion of the sample were students, and their sleep patterns may not be representative of general population. Secondly, the desired sample size of N=90 was not achieved, in part due to challenges recruiting poor sleepers as described. A post-hoc power calculation indicates that the study was therefore only sufficiently powered to detect a small effect and thus the study may have failed to identify differences between groups due to reduced power. Thirdly, the study was reliant on a sleep-diary to collate sleep parameters and while sleep-diaries are considered a gold-standard in subjective sleep assessment (Carney et al., 2012), subjective self-report can differ, with good sleepers over-estimating total sleep time and those with insomnia often under-estimating this (e.g. Harvey & Tang, 2012, Lichstein et al., 2006; Maes et al., 2014). The use of objective methods to assess sleep quality would have allowed for a) more accurate assessment of previous night’s sleep and b) scrutiny of subjective vs. objective reports for sleep quality correlated against test performance.
Future directions

There are a number of directions for future research which may help to better understand the relationship between sleep disturbance and psychotic experiences. First of all, it may be important for future studies to broaden inclusion criteria of non-clinical individuals so that factors such as anxiety and depression and their severity, can be considered alongside (degree of) sleep disturbance and psychotic-like experiences. Future studies should also aim to assess sleep using well-validated subjective and objective sleep assessment tools (see Buysse, Ancoli-Israel, Edinger, Lichsten & Morin, 2006) so that any relationships between sleep disturbance and psychotic experiences can be more accurately measured. Also, the inclusion of stress as a variable, measured psychometrically and physiologically might also facilitate an assessment interactional effects on psychotic-like experiences.

In addition, a carefully-controlled sleep restriction/ sleep deprivation study with a non-clinical sample would allow more control over sleep in determining causal relationships between impaired sleep and cognitive mechanisms underpinning psychotic (-like) experiences. Treatment outcome studies that either report on sleep outcomes following treatment for psychiatric symptoms, or report on psychiatric symptoms following treatment for sleep disturbance would also help to explore the causal mechanisms between sleep and psychiatric symptoms further.

5. Conclusion

This study sought to explore whether poor sleepers demonstrated impaired performance, relative to good sleepers, on cognitive processes that are thought to underlie hallucinations and delusions. Good sleepers and poor sleepers were not found to differ significantly across measures of psychotic-like experience or cognitive process relating to working memory, reasoning and attributional biases. A number of reasons were indicated to
offer explanation for this, largely pertaining to methodological issues including low study power, participant characteristics and sensitivity of measures. Alternatively, it could be argued that the consistent null findings indicate that poor sleep does not significantly affect the cognitive processes associated with psychotic symptoms.

It is anticipated that the findings from this study, both statistical and methodological, will inform the development of future research studies that attempt to pursue research into how sleep affects the cognitive mechanisms underpinning hallucinations and delusions.
6. References


Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L. & Morin, C. M. (2012). The consensus sleep diary:


Critical Reflection

1. Introduction

The broad aim of the research detailed within this thesis was to explore the relationship between poor sleep and psychosis. The systematic review presented in paper 1 evaluates the effectiveness of sleep treatments in serious mental illness (SMI) and the empirical study presented in paper 2 attempts to explore whether poor sleep disrupts the cognitive processes underlying psychosis. This chapter offers a reflection on the two papers detailed within the thesis and discusses key themes pertaining to methodological and theoretical issues that arose through the conduct of the studies.

2. Systematic Review

The question for the systematic review was developed to assess the current level and quality of evidence for the treatment of sleep disturbances in schizophrenia and bipolar disorder. The process of completing the review highlighted the relative dearth of empirical studies in this domain but also indicated that, of the available evidence, there seemed to be a trend towards the treatment of insomnia over any other sleep difficulty. The conclusions drawn from the review therefore suggest that insomnia can be reliably treated in patients with schizophrenia and bipolar disorder, who have existing sleep disturbances. However, as many studies demonstrated a trend towards improved sleep for their comparator groups (likely due to placebo effects which are common in sleep studies), further treatment evaluation studies with rigorous controls are required to determine the relative value of treating sleep over not treating sleep.
While the review commented on the effect of sleep treatment on sleep-related outcomes, a limitation of this review is that it does not consider broader outcomes such as the impact of treatment on psychotic experiences or quality of life. Secondary analyses examining the impact of sleep treatment on these domains could provide useful insights into the broader implications of sleep treatment. Specifically, measuring change in psychotic symptoms following sleep treatment may help to further what is known about possible causal mechanisms between sleep and psychosis. However, in the absence of consistent data (due to variation in sleep and symptom assessments) these analyses were not possible in the current review, although this would be an interesting question to pursue in subsequent reviews.

One of the strengths of this review was that it integrated evidence from both pharmacological and non-pharmacological trials and in doing so offered a unique opportunity to compare the effectiveness of a range of different interventions. Although, this presented a challenge in that it was difficult to compare the effectiveness of pharmacological interventions with non-pharmacological interventions given variation in trial methodology (e.g. differences across treatment administration) and assessment procedures (e.g. objective and subjective measures).

The process of completing the review highlighted the available controlled-trials in this domain that met the study inclusion criteria. While only two non-pharmacological trials were identified, it should be noted that a wider range of non-pharmacological approaches to sleep treatment exist, although were not represented in the review. For example, exercise (e.g. Passos, Poyares, Santana, Tufik & de Mello, 2012) and light therapy can be used to treat sleep disturbance and these interventions have been trialled in bipolar disorder (e.g. Benedetti et al., 2005; Benedetti et al., 2014). However, many of these studies involving different interventions have not used
controlled methodologies and thus were not amenable for inclusion in the review presented in paper 2. Employing less-restrictive inclusion criteria might have lead to the inclusion of many more studies that have treated sleep disturbances in the context of serious mental illness. This would have perhaps highlighted a wider range of sleep treatments and been more representative of the interventions that are available. However, in doing so might have resulted in a larger review, comprised of studies of varying design and quality. It was therefore necessary to balance number of potential studies against the potential quality of their contribution to the review. A series of scoping exercises were completed at the start of the review process and this process indicated that a sufficient number of controlled-studies could be identified that matched the proposed inclusion criteria. As such, the review proceeded with limiting studies to only those with controlled methodologies as these provide a higher quality of evidence and would therefore provide more robust findings for synthesis in the review.

The outcomes of the quality assessment however, indicated much variation in study method and quality. Firstly, there were a range of objective and subjective assessment methods which posed challenges for identifying treatment effects, particularly calculation of an effect size in relation to sleep treatment. There was also a lack of consistency in relation to the sleep parameters (e.g. sleep efficiency, total sleep time, sleep onset latency) that were collated from those studies employing objective methods of assessment. Including studies in the review which have utilised objective and subjective methodologies has highlighted a number of different approaches to sleep assessment, although simultaneously indicates a need for consensus approaches to sleep assessment (e.g. Buysse, Ancoli-Israel, Edinger, Lichstein & Morin, 2006; Carney et al., 2012).

The set of studies that were identified as part of the systematic review also varied in quality. Two studies with lowest quality ratings (Dursun et al.,
1999; Leibenluft et al., 1997 with scores of 11 and 12 respectively) were removed from the narrative synthesis for reasons relating to unclear reporting of data. The study by Shamir et al., (2000) scored only marginally higher (with a score of 13) but was retained in the review as, despite a low quality score, the study did present coherent data that could be accurately synthesised. The inclusion of the Shamir et al., study has also provided a useful contribution to the overall review in that it was one of two studies that trialled melatonin and yielded findings that were similar to Kumar et al. (2007), providing support for melatonin as a successful intervention for treating insomnia in patients with schizophrenia.

3. Empirical Study

The research presented within the thesis was developed out of a desire to explore the relationship between sleep and psychosis. While there is much evidence to suggest that lack of sleep is a common problem in mental health samples (e.g. Cohrs, 2008; Freeman, Pugh, Vorontova & Southgate, 2009; Palmese et al., 2011) and moreover, evidence to suggest that poor sleep might be related to psychotic experiences (Lee, Cho, Cho, Hang & Kim, 2012; Mulligan, Haddock, Emsley, Neil & Kyle, 2016; Reeve, Sheaves & Freeman, 2015) there had been little exploration into potential causal mechanisms. Specifically, there is a lack of clarity around whether poor sleep causes psychotic experiences (such as hallucinations) or indeed, whether it is the presence of psychotic experiences that causes poor sleep.

The study initially aimed to investigate whether sleep could contribute to psychotic-like experiences and one of the early ideas for the empirical study involved a sleep deprivation paradigm. The proposed study planned to sample good-sleeping healthy volunteers from the general population, without a history of mental health problems. Using a matched-pairs random allocation methodology, participants were intended to be
randomly assigned to sleep-as-usual (control) or sleep deprivation (experimental) conditions. The study would have taken place in carefully controlled conditions in the sleep laboratories at the University. However, given that the study’s hypothesis related to potential induction of psychotic-like symptoms, it was thought that the study would pose several ethical challenges. Such ethical challenges could be overcome with the use of rigorous safety screening processes and monitoring, however a decision was made to proceed with an alternative methodology, one in which reduced sleep was not induced but already present within the sample. The selected methodology (presented in paper 2) could achieve a similar objective but with fewer ethical challenges.

It was important that the study recruited a non-clinical sample comprising two groups whose sleep characteristics were distinct, as this would facilitate a more robust comparison. In fact, one of the strengths of the empirical study was the recruitment of two, clearly defined good and poor sleeper groups. As described in chapter 2, the PSQI scores of the poor sleep group were comparable to that of other non-psychiatric insomnia samples. The poor sleep group recruited to the empirical study also comprised individuals with low levels of anxiety and depression which reinforces the robust non-clinical characteristics of this group.

However, recruiting participants to this study, and to the poor-sleep group in particular, proved to be one of the study’s biggest challenges. Eligible participants needed to satisfy a number of criteria in order to determine whether they identified with one of the two study groups. The broad inclusion and exclusion criteria for the study are presented in Figure 1. In order to meet the minimum eligibility for the study participants needed to respond true to questions 1-4 and false to questions 5 and 6. Participants who failed to meet these criteria were excluded, to ensure that a healthy, non-clinical sample was obtained.
1. I am not currently receiving, or have recently received treatment for a mental health problem. **True/False**
2. I am not taking prescribed sleeping tablets. **True/False**
3. I am not using any psychoactive substances (substances that affect mood, perception or behaviour e.g. cannabis or other illegal drugs). **True/False**
4. I have not been in treatment for drug or alcohol related difficulties in the past 12 months. **True/False**
5. I work shifts **True/False**
6. I have a neurological or physical health problem that affects my ability to use computers. **True/False**

---

**Figure 1 Basic Eligibility Criteria for Study Inclusion**

Participants who met the basic criteria displayed in Figure 1 were then screened using a further 4 questionnaires, each with their own threshold for exclusion. It is at this stage where many participants were excluded from the study, rendering overall recruitment rates to the study quite low. The questionnaires were presented in the following order:

1. Depression Anxiety and Stress Scale (DASS-21; Lovibond & Lovibond, 1995)
2. Screening algorithm to identify other sleep disorders (Wilson et al., 2010)
3. Pittsburgh Sleep Condition Indicator (PSQI; Buysse, Reynolds, Monk, Berman & Kupfer, 1989)
4. Sleep Condition Indicator (SCI; Espie, Kyle, Hames, Gardani, Fleming & Cape, 2014)

Participants were automatically screened out if their scores on any one of the questionnaires met threshold, which meant that they did not have to unnecessarily complete all 4 questionnaires in order to be informed of the outcome. However, the electronic data were saved for every participant that attempted the screen (including those who were unsuccessful) which permitted an analysis of where participants were excluded. A graph
illustrating the number of participants that attempted the screen and those who were successful is presented in Figure 2 below.

Figure 2 Recruitment Frequency Statistics

These data were reviewed at the end of month 1 (October 2015) and were used to calculate projected recruitment over the subsequent months. Data from the first month of recruitment indicated that of the 16 people who attempted the screen only 5 were eligible, which indicated a recruitment rate of 31%. Based on this recruitment estimate, approximately 288 individuals would need to attempt the screening questions if the intended sample size of N=90 was to be achieved. Accordingly, it was agreed that the recruitment strategy needed to be widened significantly in order to advertise the study to a larger audience and increase recruitment to the study.

Additional methods of recruitment included use of social media sites Facebook and Twitter, announcements within other University faculties, Citizen Scientist, and the distribution of flyers in public places such as food outlets, hospitals and sleep-clinic waiting areas. The Student Experiment Participation Scheme (SEPS) was also used to improve recruitment to the study. The SEPS system awards Undergraduate students credits for their
participation in research studies. One credit is awarded for every 15 minutes of participation and University of Manchester students are required to obtain a certain number of credits as a requirement for the degree programme they are enrolled on. Amendments were made to the original ethics application for these additional methods of advertising.

As Figure 2 illustrates, the increased advertising appeared to have an effect on the number of volunteers attempting the screening in November and into the New Year. There was a spike in the number of participants attempting the study in February 2016 which is most likely due to the number of Undergraduate students actively seeking research studies that awarded credits prior to close of the credits system in March. The relatively low frequency of volunteers attempting the screen in December is likely due to the Christmas period resulting in fewer people around the University during the holiday period. Collectively, the use of the SEPS system to recruit students and the associated progression of recruitment (e.g. consistent with academic term) during this period emphasises the high number of students that participated in this research study. This creates some bias in the sample and reduces the representativeness of the sample and the generalizability of the study findings to the wider population.

Whilst wider advertising lead to an increase in individuals accessing the screening webpage, the limitations of some of these methods should be considered. In particular, recruitment via social media reduces control over the audiences targeted for participation. This poses challenges for understanding how representative the sample recruited is. In addition, use of the SONA Student credit system (for University of Manchester students) yielded a large proportion of student participants to the study which is not representative of the general population which was targeted for this study.

In addition to widening recruitment, the screening data were reviewed at each monthly supervision meeting to monitor the number of
participants attempting screening, including those who were subsequently screened into and out of the study (see Figure 3 below). Given the sequential screening process it was useful to review the stage at which participants were screened-out and this also helped to develop an understanding of how stringent the screening criteria were.

Recruitment progressed slowly and it was deemed necessary to revise the inclusion criteria to reduce the number of potential participants being excluded and in doing so, improve recruitment to the study. One way in which this was achieved was to revise the DASS inclusion criterion. The depression and anxiety subscales of the DASS were largely responsible for the recruitment of a non-clinical sample and as per the original criteria, individuals scoring within the moderate range were screened-out. However, given that participants were already self-reporting that they had not been in contact with mental health services for 12 months (as per basic eligibility criteria) it was thought that the DASS criterion could be relaxed somewhat without seriously compromising the integrity of the non-clinical sample. Therefore, the DASS inclusion criterion was revised to allow those participants scoring up to, (but not within) the severe range on either subscale, to proceed through the screening questions.

Although average scores for anxiety and depression remained low for both good and poor sleeper groups, the sample did include some participants with moderate levels of anxiety and depression. As such, there is less evidence to support a conclusion that low levels of anxiety and depression in the sample could account for non-significant findings. Following on from this insight, it may be useful to consider the potential relevance of anxiety and depression severity in future studies; to consider the extent to which anxiety and depression may affect performance on cognitive processes underlying hallucinations and delusions.
In light of this, it seems more reasonable to suggest that non-significant findings could be better explained through consideration of participant characteristics and other study factors, such as reduced power.

Whilst relaxing the DASS criteria improved general recruitment to the study, the number of participants eligible for the poor sleep group remained low. In order to be eligible for the poor sleep group participants originally needed to obtain a score $\geq 6$ on the PSQI and achieve a score of $\leq 16$ on the SCI. While Buysse, Reynolds, Monk, Berman & Kupfer (1989) report that the PSQI is sensitive and specific at detecting insomnia using a score of $>5$, the SCI was included as an additional criterion to ensure that a robust poor-sleep group was obtained. Accordingly, participants were originally required to satisfy two insomnia-related criterions in order to be eligible for the poor sleep group. However, this appeared to be quite restrictive and many participants who were eligible on the basis of the PSQI were screened out at the SCI stage. A decision was made to remove the SCI as a criterion and thus to permit entry to the study based on PSQI scores alone. Given that the SCI was only intended as an additional criterion, this decision to remove was not thought to seriously compromise the quality of sleep amongst the poor sleeper group. Furthermore, as described earlier/in paper 2, the PSQI scores of the poor sleeper group were in fact comparable to those of non-psychiatric insomnia groups. Each criterion change was carefully considered to maintain the rigour of the groups, whilst weighing this against the recruitment of a sample size that was sufficient for analysis.
As described, a number of procedures were trialled to increase recruitment to the study, however the projected sample size of N=90 could not be achieved within the time period for this study. As such, the study was underpowered. Starting recruitment earlier and perhaps being able to financially reimburse people might have been helpful to increasing overall sample size. However, on reflection, the recruitment of a healthy, poor sleep group might have still been a challenge given the high prevalence of anxiety and depression amongst people with sleeping problems (e.g. Carney et al., 2010; Harvey, 2002; Morphy et al., 2006, Ohayon & Roth, 2003).

The study proceeded to compare good sleepers and poor sleepers on measures that probe psychotic-like experiences. The consistent null effects for both between-group and correlational analyses indicated that either there were no effects of sleep on psychotic-experiences or that perhaps that no effects could be detected due to reduced power to detect a difference, sample characteristics (i.e. largely students) or issues pertaining to the measures used. The three questionnaires that probe proneness to hallucinations (Launay-Slade Hallucination Scale; Launary & Slade, 1981) delusions...
(Oxford Liverpool Inventory of Feelings and Experiences; Mason et al., 1995) and paranoid thoughts (Green Paranoid Thoughts Scale; Green et al., 2008) are widely used, including amongst non-clinical samples. A number of alternative measures exist however and future research may wish to consider the sensitivity of these different assessments in relation to their target sample.

The computer tasks were based on existing paradigms for working memory, reasoning and attributional biases, although, again no differences were observed between groups on any of these measures, nor were there any significant associations between previous night’s sleep and performance on these tasks. Some of the issues regarding sensitivity of the N-back and Beads Task were discussed in paper 2. However, there were some issues with the voice-detection task that are worthy of reflection here.

The voice-detection task used in the study operated from a different computer to the one where it was initially developed. This was not considered to be a potential problem until after the data had been analysed. Specifically, a low overall hit rate suggested that participants made fewer responses when compared to other studies that had used the same task (Huque, Poliakoff & Brown, 2016). This may have been due to difficulty hearing voices. An attempt was made to assess for potential difference in sound amplitude between the two computers – the one that was used to develop the programme and the one that was used in this study to test participants. A decibel meter was used to calculate the relative difference and these data are presented in Table 1.
<table>
<thead>
<tr>
<th></th>
<th>Developer Computer</th>
<th>Test Computer</th>
</tr>
</thead>
<tbody>
<tr>
<td>White noise</td>
<td>45.5</td>
<td>44.0</td>
</tr>
<tr>
<td>Loudest Voice (practice)</td>
<td>74.2</td>
<td>73.2</td>
</tr>
<tr>
<td>Loudest Voice (main task)</td>
<td>43.5</td>
<td>43.5</td>
</tr>
</tbody>
</table>

Note. Constant decibel reading of room was 42 decibels.

As illustrated in Table 1 above the decibel difference was slight, thus indicating that the amplitude of voices did not differ significantly between the two computers.

Another explanation for low hit rate could have been explained by the scoring process. As detailed in the empirical paper, the scoring of these data are based on logic and participant’s response times (obtained during the practice task) are used to generate threshold criteria for hits and false alarms in the main task accordingly. The guidance from Huque, Polikoff & Brown (2016) indicated that any responses made within 2 standard deviations of the mean response time (established during the practice) would suggest a hit and any responses made outside 2 standard deviations of the mean would normally indicate a false alarm. However, these cut-offs were considered to be a guide to data analysis and following discussion with the programme developers it was suggested that the threshold could be increased to 3. In fact, with a low false alarm rate it was sensible to see whether varying the criteria at which a voice was recognised as a false alarm (instead of a hit) would result in more false alarms. While this marginally increased the false alarm rate, these figures still did not vary significantly between groups, nor did they correlate with previous night’s sleep.
4. Summary and Conclusions

The aim of this chapter was to reflect on the findings of the study as a whole; discuss the findings in the context of the strengths and weaknesses of the two studies. Taken together the two papers presented in this thesis provide a novel contribution to the literature on sleep and severe mental illness.

Overall, the process of completing the review has been worthwhile. Despite being drawn from a relatively small number of studies, the conclusions derived from the review are nevertheless based on a robust synthesis of the available evidence. That sleeping difficulties can be reliably treated in patients with schizophrenia and bipolar disorder is an important finding; with clinical implications indicating that sleep difficulties in SMI can be treated. The review has also highlighted a number of areas for future research pursuits such as a need for more controlled studies examining the impact of sleep treatment on sleep outcomes and psychiatric symptoms and these domains recommendation for future research.

The overall finding from the empirical study was that there were no significant differences between good sleepers and poor sleepers on measures of cognitive processes in hallucinations and delusions. However, this finding should be interpreted with caution and in the context of the study’s limitations, largely that the study was under-powered and included a large student sample. While the empirical study did not yield significant outcomes it has nevertheless provided some useful methodological insights; specifically that future studies should be adequately powered and be more representative of a non-clinical general population sample. Future studies may also wish to consider the relative influence of anxiety, depression or psychosocial stressors on the relationship between poor sleep, hallucinations and delusions.
5. References


Appendices

A – Example search strategy
B – Methodological quality ratings
Appendix A

Example search strategy used in PsycInfo

1. exp Schizophrenia/
2. exp Psychosis or exp Affective Psychosis/
3. exp Schizoaffective disorder
4. exp Schizophreniform Disorder
5. exp Bipolar Disorder/
6. exp Hallucinations/ or exp Delusions
7. delusional disorder. ti,ab.
8. bipolar. ti,ab.
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. exp Sleep Deprivation/ or exp Sleep or exp Sleep disorders/
11. sleep disturb*.ti,ab.
12. exp insomnia/
13. circadian.ti,ab.
14. nightmare.mp
15. exp Hypersomnia/
16. sleep impair*.ti,ab.
17. 10 or 11 or 12 or 13 or 14 or 15 or 16
18. exp Cognitive Therapy/ or exp Cognitive Behaviour Therapy/ or exp Drug Therapy
19. exp Sleep Treatment/
20. intervention.ti,ab.
21. therap*.ti,ab.
22. medication*.ti,ab.
23. hypnotic.ti,ab.
24. exp Melatonin/
25. Chronotherapy.mp
26. 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25
27. 9 and 17 and 26
28. limit 27 to yr=“1980-2016”
### Methodological Quality Assessment Table

<table>
<thead>
<tr>
<th>Study</th>
<th>Reporting</th>
<th>External Validity</th>
<th>Internal Validity - bias</th>
<th>Internal Validity – confounding (selection bias)</th>
<th>Power</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 1 1 1 2 1 1 1 1 0</td>
<td>1 - 1</td>
<td>0 1 1 1 1 1 1 1</td>
<td>1 1 1 1 1 1 1</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>1 1 1 1 2 1 1 0 1 1</td>
<td>- - 1</td>
<td>1 0 1 1 1 1 1 1</td>
<td>- - 1 0 0 0</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>1 1 1 1 2 1 1 1 1 0</td>
<td>- - -</td>
<td>1 1 1 1 1 1 1</td>
<td>- - 1 - 1 1</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>1 1 1 1 1 1 0 1 1</td>
<td>- - 1</td>
<td>1 1 1 1 1 1 1 1</td>
<td>1 1 1 - 0 1</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>1 1 1 1 2 1 1 1 1 0</td>
<td>1 1 1</td>
<td>1 1 1 1 1 1 1 1</td>
<td>0 1 1 - 1 1</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>1 1 1 1 0 1 1 0 0 1</td>
<td>0 0 -</td>
<td>1 1 0 1 1 1 1 1</td>
<td>- - 1 - 0 0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>1 1 1 1 2 1 1 1 1 1</td>
<td>1 1 1</td>
<td>1 1 1 1 1 1 1 1</td>
<td>1 - 1 1 1 1</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>1 1 1 1 1 1 1 0 1 1</td>
<td>1 - 0</td>
<td>1 1 1 1 1 1 1 1</td>
<td>1 - 1 1 1 1</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Study</td>
<td>Item Adequately Addressed</td>
<td>Item Not Adequately Addressed</td>
<td>Item Unable to Determine</td>
<td>Total Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------------------</td>
<td>-------------------------------</td>
<td>--------------------------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yamashita et al., (2004)</td>
<td>1 1 1 1 1 1 1 1 0 1 1 1 1</td>
<td>- 1 - 1 - 1 1 1 1 1 1 0</td>
<td>- 1 0 - 0 0 0 - 0 0 1 0</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dursun et al., (1999)*</td>
<td>1 1 1 1 1 1 1 0 1 0</td>
<td>- 1 - 1 - 1 0 - 0 1 - 0 0 0</td>
<td>- 0 0 0 0 0 0</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies examining sleep in Bipolar Disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harvey et al., (2015)</td>
<td>1 1 1 1 2 1 1 1 1 1 1 1 1</td>
<td>1 - 1 1 1 1 1 1 1 1 1 1 1 1</td>
<td>0 1 1 1 1 1 1 1 0 1</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim et al., (2014)</td>
<td>1 1 1 1 1 1 1 1 1 1 1 1 1</td>
<td>1 - 1 1 0 0 1 1 1 1 1 1 1 1</td>
<td>1 - 1 0 - 0 0 0</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McElroy et al., (2010)</td>
<td>1 1 1 2 1 1 1 1 1 1 1 1 1</td>
<td>- 1 1 1 1 1 1 1 1 1 1 1 1 1</td>
<td>1 1 1 1 1 0 1 0</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norris et al., (2013)</td>
<td>1 1 1 2 0 1 1 1 1 1 1 1 1</td>
<td>1 - 1 1 1 1 1 1 1 1 1 1 1 1</td>
<td>1 1 1 1 1 0 1 0</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leibenluft et al., (1997)*</td>
<td>1 0 0 1 0 0 1 1 1 0</td>
<td>- 0 - 0 1 1 1 1 - 0 1 - 1 1 - 0 0 0 0</td>
<td></td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. 1 - item adequately addressed; 0 - item not adequately addressed; ‘-’ unable to determine. Up to 2 points can be awarded for item 5 if distributions of principal confounders are clearly described (score of 1 if this is partial). *paper not included in narrative synthesis.