Dynamic changes in sleep characteristics and pre-sleep arousal during cognitive behavioral therapy for insomnia as predictors of treatment response

by

Marissa A. Evans

Bachelor of Arts, University of Notre Dame, 2016

Master of Science, University of Pittsburgh, 2018

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2023

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Marissa A. Evans

It was defended on

April 20, 2022

and approved by

Daniel J. Buysse, Professor, Psychiatry

Thomas W. Kamarck, Professor, Psychology

Christopher E. Kline, Assistant Professor, Health and Human Development

Kathryn A. Roecklein, Associate Professor, Psychology

Aidan G. C. Wright, Associate Professor, Psychology

Dissertation Chair: Anna L. Marsland, Professor, Psychology

Dynamic changes in sleep characteristics and pre-sleep arousal during cognitive behavioral therapy for insomnia as predictors of treatment response

Marissa A. Evans, PhD

University of Pittsburgh, 2023

Abstract

Insomnia is a highly prevalent and distressing sleep disorder that is strongly associated with poor health among older adults. Cognitive-behavioral therapy for insomnia (CBT-I) is an evidence-based treatment for insomnia. Clinicians use patients' daily sleep diaries to monitor and modify treatment during each of eight sessions. Prior research on CBT-I, however, attempts to draw conclusions about the course of CBT-I by analyzing changes from pre-intervention to post-intervention only. The objective of this study was to more closely match clinical care by assessing dynamic changes in sleep characteristics and pre-sleep arousal during the course of CBT-I to predict treatment response.

Ninety-four older adults (M_{age} = 68) completed daily sleep diary data for twelve weeks (M_{days} = 70) before, during, and after CBT-I. Diary data were used to assess pre-sleep arousal, sleep latency, efficiency, quality, duration, wake after sleep onset, and alertness. We assessed dynamic changes in sleep and pre-sleep arousal by latent growth curve modeling, as well as their bidirectional association by random-intercept cross-lagged panel modeling. We used these models to predict treatment outcome, defined as the change in the Insomnia Severity Index score from pre- to post-intervention.

All models of change were best characterized by non-linear quadratic terms. We observed substantial improvements in wake after sleep onset, sleep efficiency, sleep duration, and sleep quality. We also observed statistically significant, yet less substantial, changes in pre-

sleep arousal and alertness. In models of bidirectional associations, higher levels of alertness assessed at one timepoint predicted subsequent lower levels of pre-sleep arousal at the next timepoint when adjusted for previous levels of both variables. Greater improvements in sleep efficiency and sleep quality predicted better treatment outcomes, whereas sharper initial declines in alertness portended poorer treatment outcomes. Pre-intervention data did not predict treatment outcome.

This study was the first to our knowledge to apply advanced statistical methods to model the dynamics of sleep and pre-sleep arousal data collected throughout the course of CBT-I. Our results demonstrate that documenting and observing changes during therapy has added benefit for effective patient care over contemporary pre-post intervention analyses.

Table of Contents

List of Tables

List of Figures

Preface

I would like to thank my family for supporting me in this six-year endeavor. For my husband – thank you for your unending love and for helping me to embrace my inner Viking Warrior Princess. For my Bowman family (mom, dad, Alex, Ethan, grandma) – thank you for your instrumental support, my beautiful mansion apartment, feeding me at Moon Manor, and the 24/7 tech support and car maintenance services. Acts of service is our love language. For Pippi and Heidi – I will always love you, my dearest Berners. For my granddad, Dr. Thomas E Bowman (PhD in mechanical engineering) and my dad, Dr. Paul T Bowman (PhD in chemical engineering) – thank you for showing me how to finish the PhD despite numerous challenges. I am proud to be a third-generation doctorate. For my Evans family (Mary Anne, Frank, Matt) – thank you for adopting me into your Müt. I am so pleased to be the fourth Dr. Evans at the dinner table. For my crew teammates, my book club members, and my friends – thank you for giving me an identity outside of grad school. It has been so important to my well-being to have these outlets.

I would also like to thank my academic and clinical colleagues for their support. For Dr. Martica Hall – thank you for being my academic advisor from 2016-2021. For Dr. Anna Marsland – thank you for taking over as my dissertation chair mid-project. For my dissertation committee – thank you for helping me to finish this final milestone at the intersection of my research and clinical interests. For Drs. Brant Hasler, Donna Posluszny, Scott Lewis, and Shannon Brogdon, thank you for sharing with me your passion for clinical service, and for your mentorship and guidance as I navigate clinical residency and a career as a staff psychologist.

1.0 Introduction

Insomnia is prevalent among adults over the age of 60 in the United States. Half of older adults experience symptoms of insomnia: difficulty initiating sleep, difficulty maintaining sleep, and/or early morning awakenings (Patel et al., 2018). Twelve to 20% of older adults meet criteria for insomnia disorder, which occurs when these symptoms are chronic and associated with distress or functional impairments (American Psychiatric Association, 2013a; Patel et al., 2018). The fact that over 11 million older adults experience insomnia disorder is concerning, because meta-analytic evidence has demonstrated that insomnia is associated with increased risk of developing other psychiatric disorders, cardiovascular disease, and all-cause dementia (de Almondes et al., 2016; Hertenstein et al., 2019; Sofi et al., 2014). Insomnia is also associated with increased health care utilization and higher all-cause medical costs (Wickwire et al., 2019). Thus, optimizing treatment of insomnia among older adults is critical for reducing clinical distress, improving nighttime sleep and daytime functioning, and preventing downstream negative health consequences.

Cognitive-behavioral therapy for insomnia (CBT-I) is a frontline, evidence-based treatment that improves symptoms of insomnia (Edinger et al., 2021). Understanding the process of therapeutic change during CBT-I is important for treatment refinement and improving the translation of findings from research to clinical practice (Schwartz, 2012). A systematic review identified that improvements in sleep characteristics and reductions in arousal were statistical mediators of the effect of CBT-I on insomnia symptoms (Schwartz, 2012). However, the reviewed studies only examined changes from pre-intervention to post-intervention when the patient has already left the clinic. Identifying dynamic processes that predict treatment outcomes

during CBT-I is critical, because this will allow clinicians to identify and resolve these problems when the patient is still engaged in treatment.

The current study evaluated linear and non-linear changes in sleep characteristics and pre-sleep arousal among older adults during an 8-week CBT-I intervention. These features were chosen because they are known markers of effective treatment. We assessed these changes in treatment markers in several ways: their rate and shape, their bidirectional associations, and their ability to predict improvements in insomnia symptoms.

1.1 Assessment of insomnia disorder

Individuals meet diagnostic criteria for insomnia disorder if they experience difficulty initiating sleep, difficulty maintaining sleep, and/or early morning awakenings. These symptoms must occur at least three nights per week for at least three months, and they must also be associated with clinically significant distress or impairment (American Psychiatric Association, 2013a). Importantly, insomnia disorder is not equivalent to sleep deprivation, and occurs despite the adequate opportunity for sleep (American Psychiatric Association, 2013a). Insomnia disorder is also not explained by other psychiatric disorders, medical conditions, sleep-wake disorders, or the effects of a substance (American Psychiatric Association, 2013a).

According to a clinical guideline paper by the American Academy of Sleep Medicine (AASM), the standard for insomnia disorder assessment is a clinical interview with detailed questions about sleep history, medical history, substance use history, and psychiatric history for differential diagnosis (Schutte-Rodin et al., 2008). Clinicians might choose to ask questions about the chief problem, circumstances surrounding onset, daytime consequences, current sleep-

wake schedule, the factors that are maintaining the sleep problems (in the behavioral, cognitive, work, and family/social domains), and medications that may affect sleep (Ong et al., 2021). These methods are also rated as "essential" for treatment efficacy studies, per an expert panel on the research guidelines for assessing for insomnia disorder (Buysse et al., 2006).

In the clinical guidelines, it is recommended that a measure be administered to identify comorbid disorders of sleepiness, such as the Epworth Sleepiness Scale (Johns, 1991; Schutte-Rodin et al., 2008). Polysomnography is not routinely recommended for assessing insomnia, except when there is suspicion of breathing or movement disorders (Schutte-Rodin et al., 2008). There is inconclusive evidence supporting the use of actigraphy for clinical diagnosis of insomnia Schutte-Rodin et al., 2008). In contrast to clinical guidelines, the research guidelines for treatment efficacy studies indicate that polysomnography should be used to screen all participants for comorbid sleep disorders, and either actigraphy or polysomnography should be included as outcome measures (Buysse et al., 2006).

Both clinical and research guidelines recommend that daily sleep diaries be administered for one to two weeks pre-intervention to establish sleep patterns and to rule out circadian rhythm sleep-wake disorders (Buysse et al., 2006; Ong et al., 2021; Schutte-Rodin et al., 2008). Daily sleep diaries can be used to assess several sleep characteristics each morning, including sleep latency, wake after sleep onset (WASO), sleep duration, sleep quality, and alertness. Sleep diary data have indicated that insomnia disorder is characterized by poor sleep efficiency, long sleep latency, high WASO and poor sleep quality (Gooneratne et al., 2011). Low levels of daytime alertness and short sleep duration also often occur in the context of insomnia, but are not a necessary feature of the disorder (Ancoli-Israel & Martin, 2006; cf Fortier-Brochu et al., 2012; Gooneratne et al., 2011; Hall et al., 2021).

Sleep diary data can also be used to measure pre-sleep arousal. Pre-sleep arousal is the experience of cognitive arousal (worry, ruminations) and somatic arousal (heart racing, tense muscles, upset stomach) before bedtime (Nicassio et al., 1985). Pre-sleep arousal is a specific type of anxiety, which is similarly comprised of cognitive and somatic arousal components (Byrne, 2004). However, anxiety may occur throughout the 24-hour period, whereas pre-sleep arousal is more closely linked to symptoms that occur when the individual attempts to fall asleep. Individuals with insomnia report higher levels of pre-sleep arousal compared to individuals without insomnia (Jansson-Fröjmark & Norell-Clarke, 2012; Nicassio et al., 1985). Therefore, monitoring pre-sleep arousal contributes to understanding the etiology of insomnia as well as identifying potential pathways of effective treatment.

1.2 Treatment of insomnia disorder

Insomnia disorder may be treated with cognitive-behavioral interventions, pharmacological interventions, or a combination of both. The two most common prescription medications for the treatment of insomnia are zolpidem (Ambien), a benzodiazepine receptor agonist, and trazadone, a serotonin antagonist and reuptake inhibitor (Bertisch & Buysse, 2021). Meta-analytic evidence has demonstrated that these medications have larger effects on insomnia symptoms compared to placebo control (Sateia et al., 2017). One advantage of these medications is that they have a rapid onset of action (approximately 30 minutes; Bertisch & Buysse, 2021). The most commonly used non-pharmacological intervention is CBT-I, which typically occurs in eight weekly 50-minute sessions, with maximal benefits apparent within 2-3 weeks (Morin et al., 2014). A comparative meta-analysis of behavioral treatment and pharmacotherapy suggested that

they are equally beneficial in the short-term to improve insomnia symptoms and sleep characteristics (Smith et al.,, 2002). In the long-term, CBT-I continues to show reductions in insomnia symptoms at follow-up visits occurring 12-months after treatment (Trauer et al., 2015), whereas the Food and Drug Administration has only approved the short-term use of pharmacologic treatment (4 to 5 weeks; Qaseem et al., 2016). Because pharmacologic therapy can be associated with adverse events (e.g., serious injury, fractures) and is only recommended in the short-term, the American College of Physicians recommends CBT-I as the front-line treatment for insomnia (Qaseem et al., 2016). This means that leading experts in internal medicine endorse non-pharmacological treatment - typically provided by a psychologist - for insomnia treatment.

The benefits of CBT-I instead of pharmacologic therapy are especially evident among older adults. In this population, the use of sedative-hypnotics is discouraged due to increased fall risk (Bertisch & Buysse, 2021). There is also a greater likelihood of drug-drug interactions, as older adults are more likely than younger adults to be managing comorbid medical conditions with other medications (Bertisch & Buysse, 2021). Therefore, there are even fewer reasons to select pharmacologic treatment in this population, and CBT-I is strongly recommended (Hughes & Martin, 2021).

1.3 Components of CBT-I

CBT-I is a multi-component treatment which includes psychoeducation, behavioral recommendations, and cognitive therapy. The clinician may vary the focus of treatment based on the patient's specific concerns and their daily sleep diary data. Though the timing may vary by

protocol or patient, typically the clinician introduces psychoeducation and behavioral recommendations in the first session, and cognitive therapy in the second session. In sessions 3-8, the clinician works with the patient on titration, compliance, and relapse prevention. Here, we will provide an overview of the key components occurring in sessions 1 and 2 (for more details, see CBT-I manuals; Edinger & Carney, 2015; Perlis, Jungquist, Smith, & Posner, 2005).

1.3.1 Psychoeducation about sleep

The clinician introduces models of sleep and insomnia both to provide education, as well as the rationale for the behavioral recommendations (Edinger & Carney, 2015). First, the clinician describes Borbély's two process model of sleep regulation, which states that sleep is influenced by the interrelation between the homeostatic sleep drive (Process S) and the biological clock (Process C; Borbély et al., 2016). This model suggests that sleep is impacted by both the buildup of sleep drive during the day, as well as the time of day at which an individual is trying to sleep. Optimal sleep-wake patterns occur when both sleep drive and the circadian sleep propensity are sufficiently high. According to this model, individuals with insomnia experience suboptimal sleep-wake patterns because they are attempting to sleep when their sleep drive is not sufficiently high, when their biological clock is signaling wakefulness, or both. Sleep drive increases if the individual stays awake longer, and sleep drive decreases after a daytime nap.

Next, the clinician explains Spielman's 3P model of insomnia development, which includes the predisposing, precipitating, and perpetuating factors of insomnia. Individuals are at higher risk for developing insomnia disorder if they experience predisposing factors, such as a family history of sleep disorders or a personal history of difficulty sleeping. Insomnia symptoms

6

initially develop in the context of precipitating factors, such as a stressful life event that triggers an acute episode of insomnia which lasts for several nights. Finally, the insomnia symptoms may develop into a chronic disorder that lasts for months due to perpetuating factors, such as poor sleep behaviors that maintain the symptoms of insomnia (Spielman et al., 1987).

Finally, the clinician discusses classical conditioning as it applies to insomnia. Individuals with insomnia often associate the stimulus of the "bed" with a variety of behaviors (e.g., lying in bed awake for long periods of time) and emotions (e.g., frustration, worry) that are not simply "sleep" (Perlis et al., 2021). In fact, it has been proposed that classical conditioning may be a fourth "P" of the 3P model, for "Pavlovian" (Perlis et al., 2021). This complex conditioning history, in turn, causes individuals with insomnia to experience conditioned arousal when they go to bed, which will interfere with sleep even if sleep drive is high and the biological clock is signaling sleep (Perlis et al., 2021).

The clinician also discusses healthy sleep practices, sometimes referred to as sleep hygiene. These behaviors, such as avoiding caffeine 6-8 hours before bedtime, are known to benefit sleep, but are not the active components of treatment. The active components are the behavioral recommendations and cognitive therapy.

1.3.2 Behavioral recommendations and cognitive therapy

During the first session of CBT-I, the clinician recommends three behavioral changes: time in bed restriction, stimulus control, and regularizing wake-up time. Time in bed restriction targets one of the perpetuating factors of insomnia from the 3 (or 4) P model, which is sleep extension (Perlis et al., 2021). Time in bed restriction reduces time in bed to a smaller sleep period, or "sleep window." Typically, this restriction is approximately the patient's sleep

duration plus 30 minutes, but among older adults a more modest restriction may be recommended to minimize fall risk (Hughes & Martin, 2021). Time in bed restriction increases the build-up of sleep drive (Process S) and thereby improves sleep efficiency. Once sleep efficiency has improved, the clinician may experiment with expanding the sleep window slowly to increase sleep duration. This process is a balancing act between increasing sleep duration but maintaining high sleep efficiency. The expansion of the sleep window is sometimes necessary because acute short sleep duration is associated with daytime sleepiness (Grandner et al., 2010), and chronic short sleep duration may lead to adverse physical health consequences (Itani et al., 2017).

The second behavioral change that the clinician recommends is stimulus control. Stimulus control weakens the classically conditioned association between the bed and behaviors and emotions that are not sleep. This is the "Pavlovian" P of the 4P model of insomnia (Perlis et al., 2021). Patients are instructed to only use the bed for sleep (or sex). To accomplish this, patients should only go to bed in the evening when they are sleepy, and to get out of bed if they are not asleep (Edinger & Carney, 2015). The latter instruction is sometimes altered for older adults – instead of getting out of bed, the instruction is to sit up in bed and give up the effort to fall asleep (called "counter-control" and is aimed to reduce fall risk, Hughes & Martin, 2021; Carney & Danforth, 2021). The AASM clinical guideline for CBT-I noted that the risk of nighttime falls is small, temporary, resolves during treatment, and generally tolerable (Edinger et al., 2021).

The third behavior change that the clinician recommends is to maintain a consistent wake-up time every day of the week, regardless of the previous night's sleep. This aids in setting the biological clock and, if adhered to even on a poor night's sleep, can help to build sleep drive

(Evans & Hasler, 2021). Circadian misalignment may lead to several perpetuating factors of insomnia. One may be that the insomnia symptoms are time dependent, but the patient is trying to go to bed or wake up at a different time due to work or social demands. Another may be that on free days, the patient shifts the timing of the sleep schedule and may also extend their sleep duration (Evans & Hasler, 2021).

During the second session of CBT-I, cognitive therapy techniques are introduced to help resolve both the precipitating and perpetuating factors of insomnia (Perlis et al., 2021). For example, some patients may experience excessive worry about not getting eight hours of sleep, and others may be attentionally biased to how poor their daytime functioning will be (Morin et al., 2007). Cognitive therapy techniques that have been used successfully for other disorders can reduce unhelpful beliefs about sleep, the behaviors that perpetuate these beliefs, and bedtime worry. These techniques include the identification and discussion of sleep-related automatic negative thoughts, behavioral experiments to test one's misperception of sleep, and comparing how one sleeps on nights with and without monitoring the clock (Harvey et al., 2005). Previous research has demonstrated that this combination of behavioral and cognitive therapy reduces insomnia symptoms, though there is a disconnect between how these changes are measured in the clinic compared to in the laboratory.

1.4 CBT-I outcomes

The use of daily sleep diary data is essential in clinical care, and is all but ignored in research studies. Each CBT-I session begins with reviewing the daily sleep diary collected over the past 1-2 weeks (Schutte-Rodin et al., 2008). The clinician reviews this collaboratively with

the patient to identify patterns, such as shifts in sleep timing, as well as changes in sleep duration or continuity (Evans & Hasler, 2021). Data from the sleep diary are then used to guide the treatment session. For example, if the patient has experienced a large improvement in sleep efficiency but continues to feel some decrements in alertness, then the clinician may recommend that they increase their sleep window. The use of diary data during CBT-I is consistent with measurement-based care, a patient-centered approach that uses data collected *during* treatment to inform whether patients are making progress, and if not, informs whether changes need to be made (Shimokawa et al., 2010). The importance of the sleep diary cannot be overstated, and some clinicians argue that CBT-I treatment cannot occur unless the patient has completed the diary (Perlis, 2017). However, previous CBT-I research does not match this clinical emphasis.

Most CBT-I research studies only examine changes in insomnia symptoms or sleep characteristics from pre-intervention to post-intervention. The primary outcomes are the percentage of participants who experience a reduction in insomnia symptoms (responders), and the percentage of participants who experience a reduction in symptoms to a level below a clinical threshold (remitters). A recent AASM clinical practice guideline included meta-analyses of 49 randomized controlled trials of CBT-I and reported clinically significant improvements in insomnia symptoms, response rates, and remission rates from pre-intervention to post-intervention (Edinger et al., 2021). Additional outcomes for CBT-I studies may include an evaluation of improvements in sleep characteristics and pre-sleep arousal, which have been identified as statistical mediators of treatment outcomes (Schwartz & Carney, 2012). The AASM meta-analysis demonstrated that CBT-I had a significant impact on diary-assessed sleep characteristics (latency, wake after sleep onset, efficiency, duration, sleep quality; Edinger et al., 2021). A meta-analysis of 13 intervention studes indicated that CBT-I reduced daytime

sleepiness as assessed with a self-reported retrospective questionnaire (Benz et al., 2020).

Several individual studies have shown that CBT-I reduces pre-sleep arousal (Vincent &

Lewycky, 2009; Wu et al., 2006). These previous studies are limited by only examining changes

from pre- to post-intervention, which fails to capture the dynamic therapeutic processes that

occur. This analysis is a limitation because it does not (a) identify factors for clinicians to

intervene on when the patient is still in treatment, (b) align with how clinical care occurs outside

of the laboratory, or (c) allow for non-linear change.

1.5 Change in sleep and mood during CBT-I

Only four previous studies have examined dynamic changes in sleep during CBT-I, and

only two studies have examined dynamic changes in depression and anxiety during the course of

CBT-I. These existing studies provide a clearer understanding of therapeutic processes, as well

as identifying subgroups of patients who respond optimally to treatment.

Sleep characteristics have been shown to change non-linearly during the course of CBT-I.

In a small ($n = 16$) study of time in bed restriction only (i.e., no other components of CBT-I),

subjective sleepiness increased over the first two weeks of treatment, but then subsequently

decreased, which closely matched changes in the diary-assessed sleep window (Kyle et al.,

2014). In a small ($n = 11$) study of Veterans with comorbid insomnia disorder and alcohol use

disorder, CBT-I was associated with an initial decrease, then subsequent increase in sleep

latency; there was no treatment effect on WASO or sleep duration (Chakravorty et al., 2019). In

a small ($n = 12$) study of adolescents with concussion symptoms, CBT-I was associated with

quadratic effects for sleep efficiency and sleep latency (initial improvements), a linear increase in

11

sleep duration, and no significant change in WASO (Tomfohr-Madsen et al., 2020). In the largest study to date on sleep changes during CBT-I ($n = 80$), it took two to three weeks for sleep continuity measures to reach their maximal improvement (defined as the time at which their change was nearly equivalent to the change by the end of intervention; Morin et al., 2014). Three of these studies were likely statistically underpowered to detect non-linear changes, and the study with the largest sample size did not test for non-linear changes (Morin et al., 2014). Two of these studies were conducted in samples with comorbidities (Chakravorty et al., 2019; Tomfohr-Madsen et al., 2020), which may limit the generalizability to other patient populations.

Previous studies have observed varying rates of change and varying degrees of improvement (i.e., individual differences) in anxiety and depression during CBT-I. One study of internet-based CBT-I (SHUTi) showed that anxiety and insomnia symptoms improved by the second week of treatment, and depressive symptoms improved by the fourth week, relative to a control group (Batterham et al., 2017). A second study examined CBT-I for individuals with insomnia and comorbid major depressive disorder, and demonstrated that there were three classes of responders in terms of depressive symptoms: partial-responders (68.9% of the sample), initial responders (17.6%) and optimal responders (13.5%; Bei et al., 2018). These studies did not evaluate the relationship between changes in sleep characteristics and changes in anxiety or depression across the course of treatment, though previous observational evidence suggests that they are bidirectionally associated (Winzeler et al., 2014).

Taken together, these studies indicate that there are non-linear changes in treatment markers during CBT-I, and that these changes may occur at varying rates. Only one study tested whether these changes predicted treatment outcome, which informs whether these analyses are clinically useful. However, this previous study did so by dividing participants into three

categories of change (low, medium, and high), which limited their statistical power (Morin et al., 2014). None of these studies tested whether changes in sleep characteristics and pre-sleep arousal may be associated with each other, which would improve our understanding of the time course of changes during treatment.

1.6 Overview of the current study

This study was designed to fill some of the gaps of previous studies. We assessed weekly dynamic changes in sleep characteristics and pre-sleep arousal during the course of CBT-I among older adults with insomnia. We accomplished this using statistical modeling strategies that optimized our statistical power, and in the largest sample size of sleep characteristic changes to date.

First (**Aim 1**), we tested how sleep characteristics and pre-sleep arousal changed over the intervention. We hypothesized that sleep latency and wake after sleep onset would decrease, sleep duration and alertness would initially decrease and then increase (due to time in bed restriction), and sleep quality and sleep efficiency would increase. We hypothesized that pre-sleep arousal would decrease during CBT-I.

Second (**Aim 2**), we assessed whether changes in sleep were associated with subsequent changes in pre-sleep arousal. We hypothesized that there would be a bidirectional association between sleep and pre-sleep arousal, such that better sleep would be associated with subsequent reductions in pre-sleep arousal, and conversely, that reduced pre-sleep arousal would be associated with subsequent improvements in sleep.

Third (**Aim 3**), we evaluated the extent to which baseline (intercept) and dynamic (slope) measures were associated with treatment response, defined as improvement in global insomnia symptoms. We hypothesized that, at baseline, sleep characteristics would be worse (e.g. higher sleep latency, lower sleep efficiency) and pre-sleep arousal would be higher for treatment responders. We hypothesized that, during the intervention, larger improvements in sleep characteristics and larger reductions in pre-sleep arousal would be associated with positive treatment response. The results of this aim inform whether analyzing data collected during the intervention is necessary, or whether pre-post intervention analyses are sufficient.

2.0 Methods

The current study includes data from a CBT-I study conducted within the NIH-funded program project "AgeWise: Aging Well, Sleeping Efficiently" (P01 AG20677). The intervention occurred at Western Psychiatric Institute and Clinic in Pittsburgh, PA from 2010-2015. Participants were recruited through community advertisement or referral and provided written informed consent in accordance with the Internal Review Board (IRB) at the University of Pittsburgh. Financial compensation was provided for participation. The primary aim of the program project was to identify the biological and psychological correlates of insomnia and insomnia treatments in older adults. Some of the results of the CBT-I study have been published elsewhere (Kay et al., 2015). The current study represents secondary analyses of data.

2.1 Inclusion and exclusion criteria

Participants were included if they were over the age of 60, were not experiencing cognitive difficulties (defined as a score greater than or equal to 27 on the Mini Mental Status Exam) and had no untreated psychiatric conditions or sleep disorders other than insomnia. Individuals were defined as meeting criteria for insomnia disorder if they (1) met Diagnostic and Statistical Manual-IV-Text Revision (DSM-IV-TR) criteria for primary insomnia; (2) met International Classification of Sleep Disorders (ICSD-2) criteria for general insomnia; and (3) had a score greater than 10 on the Insomnia Severity Index (ISI; Bastien et al., 2001).

During screening, participants completed 2-4 weeks of a daily sleep diary, underwent an overnight polysomnography sleep study to rule out significant sleep apnea (presence of daytime symptoms and an apnea-hypopnea index – defined as the number of breathing pauses or shallow breathing episodes per hour of sleep – greater than 20), physical examination and medical review by a study physician, and clinician-administered questionnaires. Individuals were excluded if they reported use of medications that affect sleep in the past month or use of beta blockers. They were excluded if, according to their screening daily diary, they reported consumption of more than 14 alcoholic drinks per week or more than 6 alcoholic drinks in any one day. Individuals were also excluded if the study physician deemed that they had an unstable medical condition or a condition affecting the central nervous system. Figure 1 depicts the study flow chart.

Consented to the CBT-I study
(*n* = 197)

Excluded from study (*n* = 20)
Use of medications that affect sleep (*n* = 4)
Protocol burden/too busy (*n* = 10)
Other reasons (*n* = 6)

Individuals with insomnia received CBT-I (*n* = 102)

Individuals without insomnia received no intervention (*n* = 75)

Excluded from the current study (*n* = 8)
Did not start CBT-I due to protocol burden (*n* = 5)
Discontinued from study due to medication change after first session (*n* = 1)
Discontinued from study due to non-compliance after first session (*n* = 2)

Included in this study (*n* = 94)

Not analyzed in this study

Figure 1 Study design

2.2 Study design

Baseline sleep diary data were collected for 1-2 weeks ($M = 8.9$ days, $SD = 4.2$). The first session of the intervention occurred two to four weeks after participants completed baseline diaries. The CBT-I intervention was administered in eight weekly 50-minute individual, in-person sessions. Participants completed daily diary data during the intervention for approximately 8 weeks ($M = 55$ days, $SD = 8.4$). Participants also completed daily diary data for 1-2 weeks post-intervention ($M = 6.7$ days, $SD = 3.3$), which occurred immediately after the last session. Figure 2 presents a schematic of the daily diary data collected before, during, and after the intervention.

8 continuous weeks of daily diary during intervention

1-2 weeks between

| Baseline – 1-2 wks | | 2 weeks | 1 wk | 1 wk | 1 wk | 1 wk | 1 wk | 1 wk | Post-intervention– 1-2 wks |

Visit 1   Visit 2   Visit 3   Visit 4   Visit 5   Visit 6   Visit 7   Visit 8
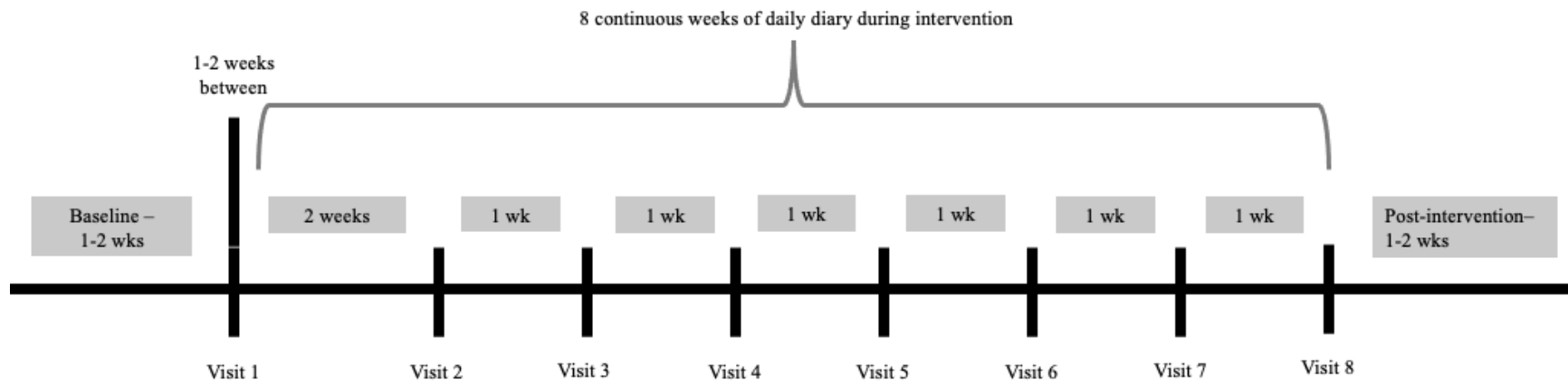
Figure 2 Daily diary assessment

The first CBT-I session included psychoeducation about sleep and an introduction to the behavioral recommendations (stimulus control, sleep restriction, and consistent wake-up time). Six specific instructions were presented: (1) select a standard wake-up time; (2) use the bed only for sleeping; (3) get up when you can't sleep; (4) don't worry, plan, etc., in bed; (5) avoid daytime napping; (6) go to bed when you are sleepy, but not before the time suggested. The second session introduced cognitive therapy components, including cognitive restructuring and constructive worry (Edinger & Carney, 2015). Sessions 3-8 included adjustments to the sleep window, encouragement, review, and reinforcement of treatment recommendations, and developing strategies to address barriers to adherence.

2.3 Measures

2.3.1 Daily diary

Participants received bound booklets of daily sleep diaries to complete each morning immediately after waking up (the Pittsburgh Sleep Diary; Monk et al., 1994). Figure 2 (on the previous page) shows the collection strategy of daily diary data. The first page of the booklet was completed as an example. During screening, study staff explained each prompt in the diary to the participant and elicited questions. Participants were instructed that if they forgot to fill out a diary, to skip that day and fill out only the current day's diary, to reduce recall bias.

2.3.1.1 Sleep characteristics

Bedtime was defined by the participant's response to: "Last night I got into bed at (__:__ AM or PM)." Wake-up time was defined as "This morning, I finally woke at (__:__ AM or PM)." These definitions of sleep timing are consistent with instructions from CBT-I to minimize time spent awake in bed. Sleep latency was defined as "I think it took me about (*response*) minutes to fall asleep." Wake after sleep onset was defined as "In total, these awakenings lasted (*response*) hours and (*response*) minutes (the TOTAL time I was awake between the time I first fell asleep and finally awakened)." Total sleep time was calculated as the number of minutes between sleep onset (bedtime plus sleep latency) and wake-up time, removing wake after sleep onset. Sleep efficiency was calculated as total sleep time divided by time in bed (wake-up time minus lights out time) multiplied by 100. If a participant's calculated total sleep time or sleep efficiency was less than zero (due to a short time in bed and long sleep latencies or wake after sleep onset), we set the value to zero.

Sleep quality and alertness upon awakening were assessed using a visual analogue scale (VAS) that was 100 mm. Scores range from 0-100. Participants used an "X" to indicate where on the VAS they would rate their sleep. For sleep quality, the VAS prompt was: "The quality of my sleep last night was" and the scale was anchored by the terms "very bad" (0) to "very good" (100). For daytime alertness, the VAS prompt was: "My alertness when I finally woke up this morning was" and the scale was anchored by the terms "very sleepy" (0) to "very alert" (100).

2.3.1.2 Pre-sleep arousal

Pre-sleep arousal was assessed daily using a VAS that was 100 mm, and therefore scores ranged from 0-100. There were three pre-sleep arousal items: "Last night at bedtime I felt mentally alert/active;" "Last night at bedtime I had trouble shutting off my thoughts;" and "Last

night at bedtime I felt worried." All three of these items had scales that were anchored by the terms "not at all" (0) to "very much" (100).

These three items were taken from the Pre-Sleep Arousal Scale, which is a 16-item measure of somatic and cognitive components of pre-sleep arousal (Nicassio et al., 1985). All three items measured cognitive components of arousal. The Pre-Sleep Arousal Scale has been shown to have high internal consistency (Cronbach's $\alpha = 0.76$) and high test-rest reliability ($\alpha = 0.72$, Nicassio et al., 1985). It has also been shown to have good discriminant validity between individuals with insomnia, poor sleepers, and normal sleepers (Jansson-Fröjmark & Norell-Clarke, 2012). The Pre-Sleep Arousal Scale is correlated at a moderate to good level with other measures assessing anxiety, depression, and dysfunctional beliefs about sleep (Jansson-Fröjmark & Norell-Clarke, 2012; Nicassio et al., 1985).

In our study, we evaluated whether our three daily diary items taken from the Pre-Sleep Arousal Scale represented one underlying factor. To do so, we first examined the correlations among the three daily diary items (averaged over the two weeks at baseline). We then examined the degree to which these three items and a composite score of the three items (average of the three items) correlated with the 16-item Pre-Sleep Arousal Scale and other measures associated with insomnia,[1] anxiety,[2] depression[3], and mental health[4] (Table 1, shown on the next page).

---

[1] Hyperarousal Scale (Pavlova et al., 2001), Ford Insomnia Response to Stress Test (Drake et al., 2004) and Dysfunctional Beliefs About Sleep-Brief Version (Morin et al., 2007)

[2] State Trait Anxiety Inventory, (Spielberger et al., 1983), Penn State Worry Questionnaire (Meyer et al., 1990)

[3] Patient Health Questionnaire-9 (Kroenke et al., 2001), Inventory of Depressive Symptomatology (Rush et al., 1986)

[4] Medical Outcomes Study, Mental Health Index (Tarlov et al., 1989)

Table 1 Correlation of daily diary items on pre-sleep arousal with other scales

	Bedtime Worry	Bedtime Thoughts	Bedtime Alert/Active	Bedtime Pre-sleep Arousal Composite
Bedtime Thoughts	0.66*			
Bedtime Alert/Active	0.21	0.29		
Bedtime Pre-Sleep Arousal Composite	0.78*	0.83*	0.69*	
Pre-Sleep Arousal, Cognitive Subscale	0.36	0.52*	0.26	0.49*
Pre-Sleep Arousal, Somatic Subscale	0.05	0.09	0.03	0.07
Pre-Sleep Arousal Scale	0.33	0.49*	0.24	0.46*
Hyperarousal Scale	0.31	0.31	0.16	0.33
Ford Insomnia Response to Stress Test	0.21	0.20	0.14	0.23
Dysfunctional Beliefs and Attitudes Towards Sleep	0.02	-0.03	-0.08	-0.04
State Trait Anxiety Inventory, Trait Scale	0.41*	0.25	-0.08	0.23
Inventory of Depressive Symptomatology	0.16	0.10	-0.19	0.01
Patient Health Questionnaire-9	0.01	-0.05	-0.22	-0.12
Medical Outcomes Study, Mental Health Index	-0.39	-0.34	0.03	-0.29

Notes. Bedtime worry is the daily diary question "Last night at bedtime I felt worried"; bedtime thoughts is the daily diary question "Last night at bedtime I had trouble shutting off my thoughts"; bedtime alert/active is the daily diary question "Last night at bedtime I felt mentally alert/active". The bedtime pre-sleep arousal composite is the average of these three items.
* indicates $p < .0001$

Table 1 (shown above) demonstrates that the daily diary items "bedtime worry" ("Last night at bedtime I felt worried") and "bedtime thoughts" ("Last night at bedtime I had trouble shutting off my thoughts") were highly correlated with each other. Both of these items were correlated with the Pre-Sleep Arousal Scale, the Hyperarousal Scale, and the Medical Outcomes Study Mental Health Index. The item on "bedtime worry" was additionally moderately correlated with State Trait Anxiety Inventory. The item on "bedtime alert/active" ("Last night at bedtime I felt mentally alert/active") was moderately correlated with "bedtime thoughts," but was not significantly correlated with any of the other scales, including the Pre-Sleep Arousal Scale.

Next, we conducted a multilevel confirmatory factor analysis, in which we tested whether the three items loaded onto a single factor at both the within-person level and at the between-person level.[5] The items on "bedtime worry" and "bedtime thoughts" had acceptable loadings at the between- and within-person levels, whereas the item on "bedtime activation" did not load strongly onto the pre-sleep arousal factor at either level (Figure 3, shown below). Based on these results, as well as the correlations reported in Table 1 (on previous page), we excluded the "bedtime alert/active" item from all future analyses and created a pre-sleep arousal factor score by calculating the average of "bedtime worry" and "bedtime thoughts." The term "pre-sleep arousal" is used to describe this variable throughout the rest of this study.
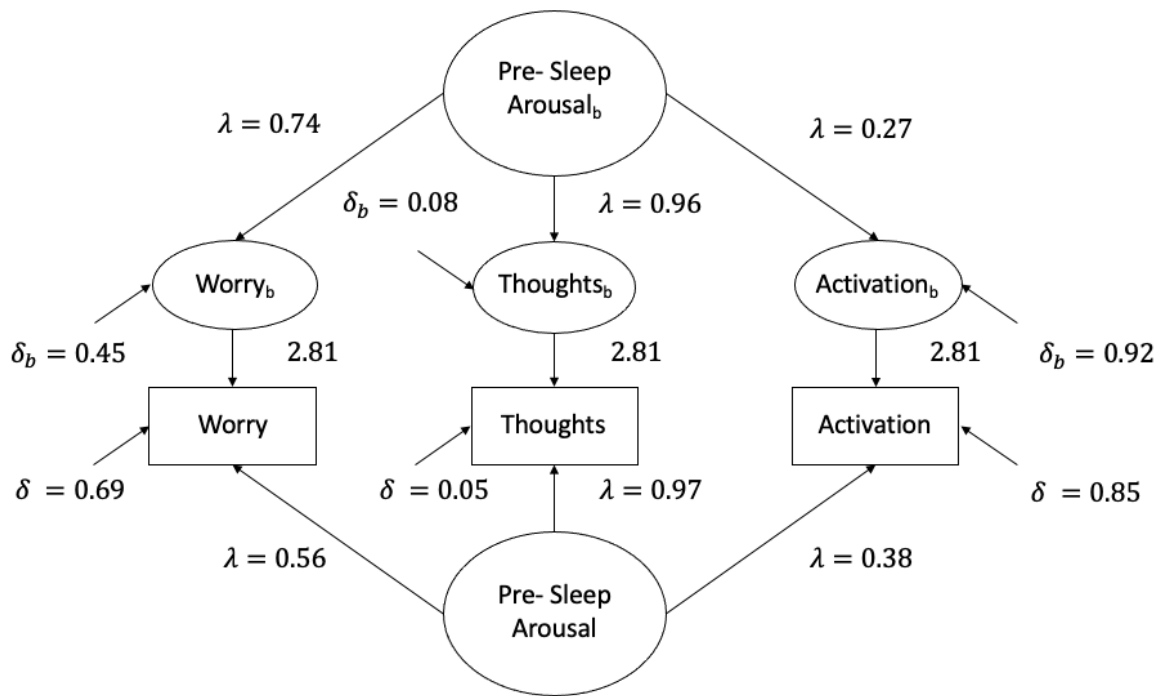
Figure 3 Multilevel confirmatory factor analysis for the pre-sleep arousal construct

---

[5] A factor loading cut-off of 0.40 was used, which indicates that 16% of the variance in the item is explained by the factor (Yong & Pearce, 2013).

2.3.2 Insomnia Severity Index

The Insomnia Severity Index (ISI) evaluates insomnia symptoms and associated clinical distress using seven items; each item can be rated on a scale of 0 to 4, so that the total score ranges from 0 to 28, with higher values indicating more severe insomnia (Bastien et al., 2001). This measure has been demonstrated to be internally consistent (Cronbach's alpha = 0.74), and to have high concurrent validity compared to clinician ratings in a sample of older adults (Bastien et al., 2001).

The ISI was administered at baseline and post-intervention. We created a simple change score variable for the ISI, calculated as the score at baseline minus the score at post-intervention. This calculation means that higher change score values indicate a greater reduction in insomnia symptoms. Insomnia responders were defined as participants whose ISI change score was greater than or equal to an 8-point decrease. Insomnia remitters were defined as participants whose ISI score at post-intervention was less than 8 (indicating "no significant insomnia"). These cut-offs are consistent with those of the AASM clinical practice guideline (Edinger et al., 2021).

2.4 Statistical analysis plan

Ninety-four participants provided data at baseline and during the intervention, and 64 participants provided post-intervention data. We calculated the average values of sleep characteristics and pre-sleep arousal separately at baseline, during the intervention, and post-

intervention. This means that, for example, each participant had one value for sleep efficiency at baseline, one value for sleep efficiency during the intervention, and one value for sleep efficiency at post-intervention. We used these data to examine the between-person correlations (Table S1; Supplemental Tables are shown in Section 5, beginning on page 56) and within-person correlations[6] (Table S2) among sleep characteristics and pre-sleep arousal. The goal of these analyses was to determine whether the two diary-assessed constructs were separable despite a common measurement method. As shown in Table S1, at baseline, pre-sleep arousal had a very small correlation with sleep measures ($r = -0.01$ to 0.04), whereas many of the sleep variables were very strongly correlated with each other ($r = -0.88$ to 0.71). This suggests that the pre-sleep arousal construct is separable from the sleep construct. The largest correlations among the sleep variables were among WASO, sleep latency, and sleep efficiency, which is unsurprising given that sleep efficiency is calculated using the other variables. Additionally, sleep quality and sleep efficiency were highly correlated ($r = 0.65$).

We assessed the intraclass correlation coefficient for each diary-assessed variable across the three time points to determine the degree of within-person variance in our study (Table S3). For sleep latency, sleep efficiency, and sleep duration, the within-person variance was greater than 50% across all time points. Sleep quality and pre-sleep arousal were above 50% at baseline and during the intervention but were below this threshold at post-intervention. Alertness had the least amount of within-person variance, with 47.6% at baseline, 38% during the intervention, and 26.3% post-intervention.

---

[6] Repeated measures correlations account for non-independence by removing measured variance between-participants, and assesses the overall intra-individual association between two measures (Bakdash & Marusich, 2017; Bakdash & Marusich, 2021).

We tested a series of structural equation models (SEM) to address our three aims of evaluating dynamic changes in sleep and pre-sleep arousal during CBT-I. We selected SEM for several reasons: SEM does not assume that predictors are measured without error, models can accommodate multiple simultaneous outcomes, slopes can be predictors, and lagged effects can be tested in both directions in one model (Sadikaj et al., 2019).

Diary data were averaged across six timepoints: baseline, CBT-I sessions 1 and 2, CBT-I sessions 3 and 4, CBT-I sessions 5 and 6, CBT-I sessions 7 and 8, and post-intervention. Statistical models have been previously shown to be increasingly computationally demanding after eight time points (Preacher et al., 2012), and preliminary examination of models with 10 timepoints (for weekly CBT-I sessions) resulted in a high proportion of non-convergent models.

2.4.1 Modeling dynamic changes in sleep and pre-sleep arousal

We used latent growth curve models to assess the degree to which sleep characteristics and pre-sleep arousal changed over time (**Aim 1**). This analysis allowed us to understand both the shape of change and the rate of change of these variables during the intervention. If a linear model is best, it suggests that previous pre-to-post intervention analyses are sufficient for capturing change. If, however, another model is preferable, then the results support the use of data collected during the intervention in future CBT-I studies.

First, we tested the inclusion of a linear slope term, and then we evaluated the inclusion of quadratic and cubic terms. Quadratic and cubic terms can improve model fit over the linear model with an increasing number of timepoints (Preacher et al., 2012). A quadratic term can represent a U-shaped association, or it can also capture the "hockey stick effect," which is an initial improvement in symptoms, followed by a subsequent stability in symptoms or slight

decrement in symptoms (Hoberg et al., 2017). Additionally, we evaluated a three-piece piecewise function (baseline, intervention, post-intervention), because a previous study of depressive symptom changes during CBT-I demonstrated that piecewise functions were needed to account for different slopes during baseline compared to during therapy (Bei et al., 2018).

Testing of latent growth curve models indicated that the cubic models and piece-wise models did not converge. Linear and quadratic models did converge. We compared linear versus quadratic models for the sleep characteristics and pre-sleep arousal using the chi-squared statistic (χ^2), Akaike's Information Criteria (AIC), Bayesian Information Criteria (BIC), the Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMSR) (Table S4). For all variables, the quadratic model was a better fit of the data than the linear model. This result suggests that there are non-linear changes occurring during CBT-I, and that previous pre-to-post intervention analyses do not adequately capture this effect.

Though the quadratic model was better relative to the linear model, it does not necessarily mean that it is an acceptable model of the data; we assessed whether it was acceptable using the CFI, TLI, RMSEA, and SRMSR. We used the following cut-off values: CFI or TLI $\geq$ 0.95, RMSEA $\leq$ 0.06, and an SRMR $\leq$ .08 (Hu & Bentler, 2009; Shi et al., 2019).[7] For all models with acceptable fit, we evaluated the following parameters: the mean intercept, which represents the average of the variable at baseline; the mean slope (linear term), which represents the average rate of change of the variable; the mean quadratic, which represents the average shape of the non-linear change; the intercept-slope covariance, which represents the correlation between the baseline value and rate of change; the intercept-quadratic covariance, which

---

[7] These cut-off values were be used throughout the methods and results to determine acceptable model fit.

represents the correlation between the baseline value and the shape of the quadratic; and the slope-quadratic covariance, which represents the correlation between the rate of change and the shape of the quadratic (Preacher et al., 2012).

After evaluating the sleep characteristics in separate quadratic models, we built a series of multivariate latent growth curve models to evaluate the extent of one sleep variable's change accounting for the other sleep variables in that same model. This analysis allows for an evaluation of whether change in one variable predicts change in another variable. Because of the high correlations between WASO, sleep latency, and sleep efficiency, as well as for sleep efficiency and sleep quality (Table S1), we did not create multivariate latent curve models with these variables entered concurrently, as they would be highly collinear.

2.4.2 Modeling the bidirectional associations between sleep and pre-sleep arousal

Next, we built a series of models to evaluate the bidirectional and cross-lagged associations between pre-sleep arousal and each sleep variable (**Aim 2**). This analysis allows us to understand the dynamic association between two markers of treatment to improve our understanding of why the treatment works.

We examined each sleep characteristic in association with pre-sleep arousal separately, instead of creating a latent factor for sleep, because we *a priori* hypothesized that some sleep characteristics may be more strongly associated with pre-sleep arousal than others. Though there were small between-person correlations between sleep characteristics and pre-sleep arousal at baseline (Table S1), these correlations do not inform us of whether or not the change in one variable predicts the change in another variable.

We included both autoregressive and cross-lagged terms between the residuals of sleep and pre-sleep arousal to examine timepoint-to-timepoint associations between two variables (Curran et al., 2014). These terms were created by regressing the residual at time t on the residual at time $t-1$ across the two constructs (Curran et al., 214). The autoregressive component accounts for stability from timepoint-to-timepoint, and the cross-lagged component allows for associations between sleep at one timepoint to predict pre-sleep arousal at the next timepoint (and vice versa).

First, we attempted to model these associations using multivariate latent growth curve models with structured residuals. These models included the intercepts, slopes, and quadratic terms from the latent growth curve models identified in **Aim 1**, plus an autoregressive component and cross-lagged component to test the bi-directional associations between one sleep characteristic and pre-sleep arousal (Mund & Nestler, 2019). The multivariate latent growth curve models with structured residuals did not converge for any of the sleep variables.

Second, we modeled these associations using cross-lagged panel models. These cross-lagged panel models include the autoregressive and cross-lagged terms, but do not include the growth modeling structure identified in **Aim 1**. The first method we attempted was a random-intercept cross-lagged panel model with constrained autoregressive and cross-lagged terms. The random intercept allows for individual differences in the starting points of sleep and pre-sleep arousal (Hamaker et al., 2015). The constraints assume equivalence across timepoints for the autoregression terms and the cross-lagged terms. For example, sleep latency at the first timepoint predicting sleep latency at the second time point is constrained to be equivalent to sleep latency at the fourth time point predicting sleep latency at the fifth time point. The random-intercept cross-lagged model with constrained terms converged.

We compared the random-intercept model to a traditional cross-lagged panel model. For all variables, chi-squared difference tests indicated that the inclusion of a random-intercept fit the data better than no inclusion of a random intercept. This suggests that individual differences in baseline variables are important to include in the model.

Third, we allowed for the random-intercept cross-lagged panel model to have unconstrained autoregressive and cross-lagged terms. Unconstrained terms allow for the regressions to vary at each time point (Curran et al., 2014). In other words, comparing this model to the constrained model allows us to test whether the strength of the association between a sleep characteristic and pre-sleep arousal differed across the course of the intervention. The models did not converge when we allowed for unconstrained terms. Thus, we determined that the random-intercept cross-lagged panel model with constraints was the optimal model building strategy for **Aim 2**. Figure 3 (shown on the next page) provides a visual depiction of this modeling strategy.
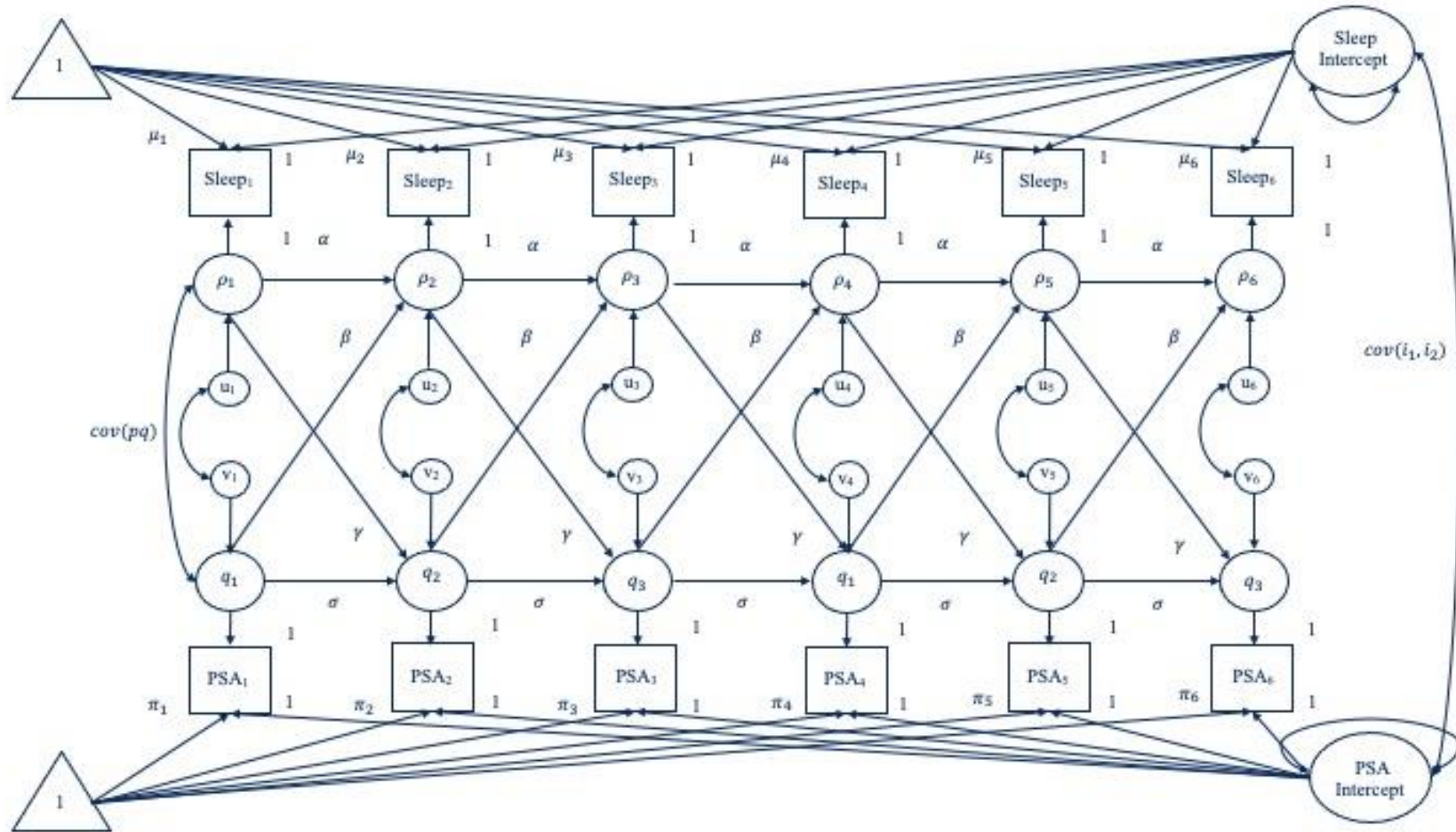
Figure 4 Random intercept cross-lagged panel model for sleep and pre-sleep arousal.

Notes. PSA, pre-sleep arousal; μ, group mean sleep for that timepoint; α, autoregression term for sleep; ρ, latent variable of sleep; β, cross-lagged term of PSA predicting sleep; u, error term for sleep; v, error term for PSA; γ, cross cross-lagged term of sleep predicting PSA; q, latent variable of PSA; σ, auturegression term for PSA; π, group mean PSA for that timepoint; $cov(pq)$, covariance of sleep latent variable and PSA latent variable; $cov(pq)$, covariance of sleep intercept and PSA intercept

32

We evaluated the fit indices to determine which models had acceptable model fit. For models with acceptable fit, we evaluated the following terms: sleep autoregressive term, pre-sleep arousal autoregressive term, cross-lagged term of sleep predicting pre-sleep arousal, and the cross-lagged term of pre-sleep arousal predicting sleep.

2.4.3 Modeling treatment response in association with dynamic changes

Finally, we extended the latent growth curve models to test whether changes in sleep characteristics and pre-sleep arousal were predictive of treatment response (**Aim 3**). This analysis informs whether modeling dynamic changes during the intervention ultimately predict improvement in insomnia symptoms. In other words, it addresses whether or not these analyses are clinically relevant, and informative beyond pre-post intervention analyses.

In terms of treatment outcomes in our study, 63 participants were both responders and remitters, 6 participants were responders but not remitters, 6 participants were remitters but not responders, and 17 participants were neither responders nor remitters. Two participants did not complete the ISI at post-intervention. Given the limited number of participants in the non-responder or non-remitter categories ($n = 23$) and the complexity of our models, we chose not to use a multiple group modeling strategy, which would stratify results based on whether a participant was a responder or non-responder (Lefcheck, 2021). Instead, we defined treatment response as the continuous ISI change score (ISI at post-intervention minus ISI at baseline).

We initially attempted to create models with treatment response as the outcome. None of the latent growth curve models. Next, we developed models with treatment response as a between-person predictor (Mulder & Hamaker, 2021). All models converged. Therefore, we chose to include treatment response as a predictor to test **Aim 3**.

Consistent with **Aim 1** and **Aim 2**, we evaluated the fit indices for the latent growth curve models with the inclusion of the treatment outcome term. For latent growth curves with acceptable fits, we evaluated whether treatment outcome was significantly associated with baseline, rate of change, or shape of change of sleep characteristics or pre-sleep arousal.

2.4.4 Summary of statistical models

For **Aim 1**, we conducted six univariate latent growth curve models and five multivariate latent growth curve models. For **Aim 2,** we conducted six cross-lagged panel models. For **Aim 3**, we constructed six latent growth curve models. Thus, we built 23 statistical models total, each with several regression slopes, intercepts, and covariances. Given the number of tests, we adjusted our alpha threshold to 0.01 (Smith & Cribbie, 2013). All of our tables indicate with an asterisk (*) if the effect meets this criterion. We only present and interpret in-text the effect sizes that had an associated p-value that was below 0.01. and the focus is on the strength of the effect (small, medium, or large).

Analyses were conducted in R version 4.1.0 (R Core Team, 2021). We prepared the data for analysis using the following packages: "tidyr" (Wickham, 2021), "plyr" (Wickham, 2011), "dplyr" (Wickham, Francois, & Muller, 2021), "Hmisc" (Harrell, 2021), "Rmisc" (Hope, 2013), and "lubridate" (Grolemund & Wickham, 2011). We evaluated descriptive statistics using the package "psych" (Revelle, 2021) and rmcorr (Bakdash & Marusich, 2021). We used the packages "ggplot2" (Wickham, 2016) and "gridExtra" (Auguie, 2017) for visualizing the data. We conducted our SEM models using "lavaan: (Rosseel et al., 2021), and printed our tables using the packages "stargazer" (Hlavek, 2018), "semTools" (Jorgensen, Pornprasertmanit,

Schoemann, & Rosseel, 2021), and "semTable" (Johnson & Kite, 2020). We used the syntax for random-intercept cross-lagged panel models from Flourney (2020).

3.0 Results

Table 2 (on the next page) includes the descriptive statistics of our study at baseline, during the intervention, and post-intervention. At baseline, participants were 68 years of age on average, 70% of the sample was female, and the average score on the Insomnia Severity Index was approximately 17, which indicates moderate insomnia symptoms. Before the intervention, the average participant slept 6.5 hours, took longer than 30 minutes to fall asleep, had greater than 60 minutes of wakefulness after sleep onset, and a sleep efficiency of 79.4%. On a scale of 0-100, with higher values indicating greater arousal, the average participant had a score of 31.5 on the pre-sleep arousal variable. During the intervention, the average participant slept 6 hours, took 19 minutes to fall asleep, had 35 minutes of wakefulness after sleep onset, a sleep efficiency of 87.1%, and a pre-sleep arousal score of 25.

Post-intervention data were available for 64 participants. These participants did not differ from the full sample on any baseline characteristics, including insomnia severity. The average participant at post-intervention had an Insomnia Severity Index score of 5, which indicates the absence of insomnia symptoms. After the intervention, the average participant slept 6.3 hours, took 19 minutes to fall asleep, had 32 minutes of wakefulness after sleep onset, a sleep efficiency of 88.9%, a sleep quality score of 50 (scale of 1-100), an alertness rating of 65.4 (scale of 1-100), and a pre-sleep arousal score of 22.4 (scale of 1-100).

Table 2 Descriptive statistics

	Baseline M(SD), min-max, n = 94	Intervention M(SD), min-max, n = 94	Post-intervention M(SD), min-max, n = 64
Age, years	68.0 (7.2), 60-93	68.0 (7.2), 60-93	68.3 (7), 60-90
Sex, female n (%)	66 (70%)	66 (70%)	49 (77%)
Insomnia Severity Index, 0-28	16.9 (3.7), 11-28	Not assessed	5.4 (4.8), 0-19
Diary data			
Number of days	8.9 (4.2), 5-38	54.5 (8.4), 32-77	6.7 (3.3), 1-14
Sleep latency, min	32.4 (30.7), 5-180	19.1 (22.0), 2-199	18.7 (29.4), 2-229
WASO, min	69.8 (47.7), 0-250	35.5 (32.0), 4-193	31.8 (32.7), 0-201
Sleep efficiency, %	79.4 (11.7), 39-99	87.1 (9.7), 40-98	88.9 (9.0), 43-99
Sleep duration, hours	6.5 (1.1), 2.8-8.9	6.0 (0.9), 3.4-8.3	6.3 (0.9), 4.3-8.5
Sleep quality, 0-100	50.0 (16.3), 4-82	63.6 (16.2), 9-89	69.3 (19.3), 19-95
Alertness, 0-100	59.3 (19.5), 3-94	60.0 (22.3), 7-96	65.4 (25.4), 9-97
Pre-Sleep Arousal, 0-100	31.5 (14.8), 3-72	24.6 (14.2), 3-63	22.4 (17.1), 2-69

Notes. WASO, wake after sleep onset. Higher values of sleep quality and alertness indicate better values, higher values of pre-sleep arousal indicate greater arousal.

3.1 Dynamic changes in sleep and pre-sleep arousal

Our **Aim 1** was to assess how sleep and pre-sleep arousal variables dynamically changed (in both rate and shape) over the course of the CBT-I intervention. Figure 5 (on the next page) shows the pattern of the change in these variables across the intervention.
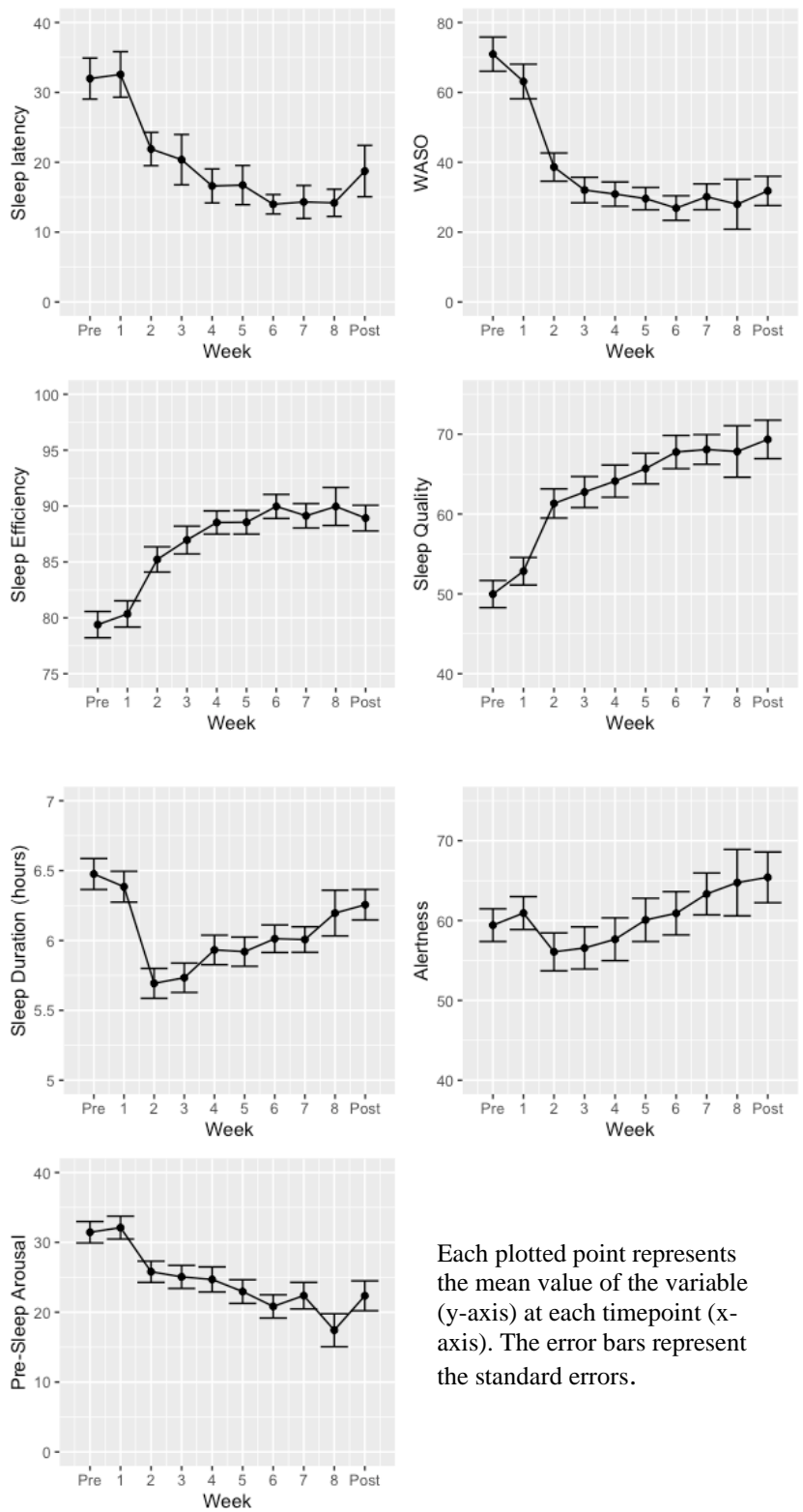
Each plotted point represents the mean value of the variable (y-axis) at each timepoint (x-axis). The error bars represent the standard errors.

Figure 5 Weekly sleep diary data.

We evaluated the model fit indices for all seven of our quadratic latent growth curve models (Table S4). Based on one or more of our cut-off criteria[8], all of our latent growth curve models had acceptable model fit, except for sleep onset latency. Because the sleep onset latency data could not be adequately modeled, we do not present or interpret the latent variables for this model.

Sleep quality, sleep efficiency, pre-sleep arousal, and WASO all demonstrated initial steep improvements in characteristics, followed by a leveling off of change, consistent with the "hockey stick effect." Alertness initially worsened, then improved steadily, and was higher by the end of treatment than it was pre-intervention. The change in sleep duration resembled more of a U-shaped association, characterized by an initial decrease, followed by a steady increase (and no leveling off/asymptote).

Table 3 (on the next page) shows that the magnitude of the linear and quadratic terms were large[9] for wake after sleep onset (WASO), sleep efficiency, sleep duration, and sleep quality (β[10] ranged from |0.85-1.31|). The magnitude in change was medium-sized for pre-sleep arousal (β_{slope} = -0.69, $\beta_{quadratic}$= 0.49), and small for alertness (β_{slope} = -0.15, $\beta_{quadratic}$= 0.46).

---

[8] As mentioned on page 27, we used the following cut-off values: CFI or TLI $\geq$ 0.95, RMSEA $\leq$ 0.06, and an SRMR $\leq$ .08.

[9] Small effect size is 0-0.29, medium is 0.3-0.49, large is > 0.5.

[10] β indicates that the estimate was standardized, such that the observed and latent variances are equal to 1. This is similar conceptually to a standardized beta coefficient in regression (Scheidel, 2019).

Table 3 Dynamic changes in sleep and pre-sleep arousal during CBT-I

	WASO	Sleep efficiency	Sleep duration	Alertness	Sleep quality	Pre-sleep arousal
Latent Intercepts						
Intercept, B(SE)	71.12(4.77)*	78.95(1.16)*	385.55(6.26)*	59.22(1.99)*	49.82(1.64)*	31.98(1.47)*
Linear Slope, Beta(SE)	-1.31(0.20)*	1.25(0.18)*	-0.85(0.16)*	-0.15(0.14)	1.09(0.18)*	-0.69(0.16)*
Quadratic Slope, Beta(SE)	1.11(0.20)*	-0.97(0.18)*	0.91(0.19)*	0.46(0.18)*	-0.80(0.17)*	0.49(0.17)*
Latent Covariances, Beta(SE)						
Intercept w/Linear Slope	-0.72(0.07)*	-0.42(0.11)*	-0.40(0.11)*	0.06(0.15)	-0.24(0.13)	-0.35(0.13)*
Intercept w/Quadratic Slope	0.63(0.10)*	0.30(0.13)	0.31(0.14)	-0.14(0.19)	0.18(0.14)	-0.36(0.14)*
Linear Slope w/Quadratic Slope	-0.96(0.01)*	-0.96(0.01)*	-0.99(0.01)*	-0.94(0.02)	-0.94(0.02)*	-0.93(0.03)*

Notes. WASO, wake after sleep onset; SE, sleep efficiency; TST, total sleep time; PSA, pre-sleep arousal; B, unstandardized coefficients; SE, standard error; Beta, standardized coefficient; * $p < .01$

The standardized latent covariances in Table 3 (on the previous page) represent the correlation between intercepts and slopes. There were several baseline variables that predicted the rate and shape of change in these variables during the intervention. Individuals who had higher sleep efficiency ($r = $ -0.42) or longer sleep duration ($r = $ -0.40) at baseline experienced a slower change in these variables during the intervention. Individuals who had higher levels of WASO at baseline experienced a slower initial change ($r = $ -0.72), and subsequent faster leveling off ($r = $ 0.63) of their WASO, than individuals who began the intervention with lower levels of WASO. Individuals with a greater level of pre-sleep arousal at baseline demonstrated slower initial changes ($r = $ -0.35), but later experienced faster improvements ($r = $ -0.36) in pre-sleep arousal relative to those with lower levels of pre-sleep arousal. Baseline values of alertness and sleep quality were not associated with the rate of change of these variables during the intervention ($r = $ 0.06 and -0.24, respectively).

As shown in Table 3 (on the previous page), all of our models had a large negative covariance between the linear slope and the quadratic slope ($rs = $ -0.93 to -0.99). For sleep efficiency, sleep quality, WASO, and pre-sleep arousal, this suggests that faster improvements in the variable were associated with slower leveling off of change. For sleep duration and alertness, this means that a greater decline in these variables initially was associated with a slower improvement in these variables later in the intervention.

Figure 6 (on the next page) extends the visual depiction of the data shown in Figure 5 by including an overlay of these latent growth curve models, indicated by the blue line.
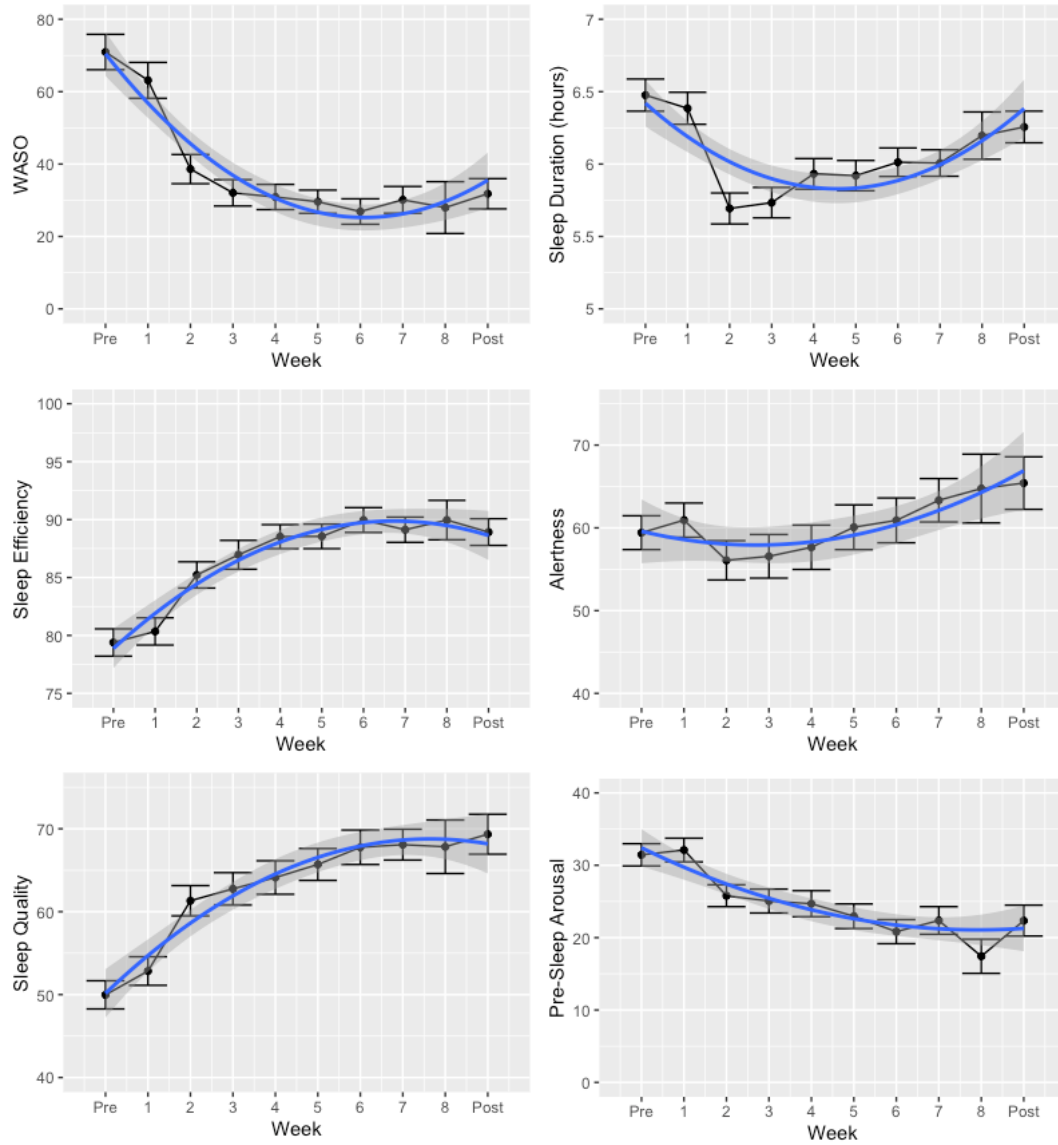
Figure 6 Weekly sleep diary data with latent growth curve models

Each plotted point represents the mean value of the variable (y-axis) at each timepoint (x-axis). The error bars represent the standard errors. The blue line depicts the latent growth curve model that was fit to the data.

3.1.1 Dynamic changes in multiple sleep variables concurrently

Next, we built a series of multivariate latent growth curve models that included sleep variables modeled concurrently as outcomes. We were able to model two sleep variables concurrently, but none of the models with three sleep variables converged. The fit indices are shown in Table S5. We present the standardized estimates for acceptable models in Table 4 (on the next page). The purpose of this analysis is to understand whether changes in one sleep variable are associated with changes in other sleep variables during CBT-I.

We found that better sleep quality at baseline was associated with faster improvements (β = 0.47) and slower leveling off of change for WASO (β = -0.42), while faster improvements in sleep quality during CBT-I were associated with slower improvements (β = -0.56) and faster leveling off of change for WASO (β = 0.57). While there were correlations in changes between sleep efficiency and alertness, as well as sleep duration and alertness, we did not interpret these given that the slopes for alertness and duration were non-significant.

Table 4 Dynamic changes in two sleep variables concurrently during CBT-I

	WASO + Sleep quality	Sleep efficiency + Alertness	Sleep duration + Alertness	Sleep duration + Sleep quality
Latent Intercepts				
Intercept, V1, B(SE)	71.16(4.77)*	78.95(1.16)*	385.56(6.26)*	385.52(6.27)[*]
Intercept, V2, B(SE)	49.96(1.64)*	59.25(1.99)*	59.25(1.99)*	49.76(1.64)[*]
Linear Slope, V1, Beta(SE)	1.08(0.17)*	1.25(0.18)*	-0.16(0.14)	1.08(0.17)*
Linear Slope, V2, Beta(SE)	-1.34(0.2)*	-0.15(0.14)	-0.84(0.16)*	-0.86(0.16)*
Quadratic Slope, V1, Beta(SE)	1.14(0.2)*	-0.97(0.18)*	0.91(0.18)*	0.93(0.19)*
Quadratic Slope, V2, Beta(SE)	-0.79(0.17)*	0.46(0.18)*	0.47(0.18)*	-0.78(0.17)*
Latent Covariances, Beta(SE)				
Intercept V1 w/Intercept V2	-0.66(0.07)*	0.14(0.11)	0.03(0.12)	0.45(0.1)*
Intercept V1 w/Linear Slope V2	0.09(0.13)	-0.02(0.14)	0.07(0.14)	0.21(0.13)
Intercept V1 w/Linear Slope V1	-0.72(0.07)*	-0.41(0.11)*	-0.40(0.11)*	-0.40(0.11)*
Intercept V1 w/Quadratic V1	0.63(0.1)*	-0.01(0.17)	0.31(0.13)	0.3(0.14)
Intercept V1 w/Quadratic V2	-0.03(0.14)	0.3(0.13)	-0.02(0.17)	-0.19(0.15)
Intercept V2 w/Linear Slope V2	-0.23(0.13)	0.07(0.15)	0.07(0.15)	-0.23(0.13)
Intercept V2 w/Linear Slope V1	0.47(0.12)*	0.04(0.14)	0.14(0.14)	0.02(0.14)
Intercept V2 w/Quadratic V1	-0.42(0.13)*	-0.16(0.19)	-0.17(0.15)	-0.06(0.15)
Intercept V2 w/Quadratic V2	0.16(0.14)	-0.01(0.15)	-0.17(0.19)	0.16(0.14)
Linear Slope V1 w/Linear Slope V2	-0.56(0.13)*	0.39(0.15)*	0.35(0.16)	0.28(0.15)
Linear Slope V2 w/Quadratic V1	0.57(0.14)*	-0.94(0.02)*	-0.38(0.17)	-0.33(0.16)
Linear Slope V2 w/Quadratic V2	-0.94(0.02)*	-0.39(0.16)*	-0.94(0.03)*	-0.94(0.02)*
Linear Slope V1 w/Quadratic V1	-0.96(0.01)*	-0.57(0.18)*	-0.99(0.01)*	-0.98(0.01)*
Linear Slope V1 w/Quadratic V2	0.58(0.14)*	-0.96(0.01)*	-0.56(0.18)*	-0.35(0.16)
Quadratic V1 w/Quadratic V2	-0.66(0.14)*	0.59(0.18)	0.61(0.19)*	0.42(0.17)*

Notes. WASO, wake after sleep onset; B, unstandardized coefficients; SE, standard error; Beta, standardized coefficient; V1, first variable in the column name; V2, second variable in the column name;* p <.01

3.2 Bi-directional cross-lagged associations between sleep characteristics and pre-sleep arousal during CBT-I

Our **Aim 2** was to evaluate whether dynamic changes in sleep characteristics and dynamic changes in pre-sleep arousal were associated with each other over the course of CBT-I. This analysis allows us to understand the relations between two markers of treatment to improve our understanding of how therapeutic change unfolds. Figure 4 (on page 32) provides a visual depiction of the random intercept cross-lagged panel model with constrained terms (i.e., autoregression term is equivalent across time, cross-lagged term is equivalent across time).

We evaluated the model fits for the random-intercept cross-lagged panel model for the association between pre-sleep arousal and the sleep characteristic variables (Table S6). We present and interpret the results of the models with acceptable model fit in Table 5 (on the next page).

Table 5 Bidirectional associations between sleep and pre-sleep arousal during CBT-I

	WASO	Sleep efficiency	Sleep quality	Alertness
Regression slopes, Beta (SE)				
Sleep autoregressive (α)	0.36(0.07)*	0.34(0.06)*	0.38(0.06)*	0.77(0.05)*
PSA autoregressive (σ)	0.75(0.06)*	0.44(0.08)*	0.46(0.09)*	0.43(0.04)*
Sleep predicting pre-sleep arousal (γ)	-0.10(0.08)	-0.01(0.08)	-0.02(0.05)	-0.27(0.08)*
Pre-sleep arousal predicting sleep (β)	-0.09(0.14)	-0.13(0.06)	-0.05(0.06)	-0.05(0.07)
Latent Covariances, Beta (SE)				
Sleep Intercept w/PSA Intercept $cov(i_1, i_2)$	0.01(0.22)	0.01(0.12)	-0.26(0.12)	-0.04(0.18)
Sleep Latent Variable w/PSA Latent Variable $cov(pq)$	0.01(0.15)	-0.04(0.12)	0.05(0.12)	0.28(0.29)

Notes. WASO, wake after sleep onset; Beta, standardized coefficient; SE, standard error; the symbols in the first column (e.g., β) match the symbols in Figure 3. * indicates p < .01

Only one cross-lagged association was significant (Table 5, shown above). Higher levels of alertness predicted subsequent lower levels of pre-sleep arousal (β = -0.27), adjusting for previous periods' alertness and arousal. Additionally, these two variables had the largest autoregressive terms, which suggests strong relationships from one time point to the next. This could indicate higher stability, or it could indicate more linear change.

3.3 Associations between dynamic intervention changes and treatment response

Lastly, for **Aim 3**, we assessed whether treatment response was associated with dynamic changes in sleep characteristics or pre-sleep arousal. This analysis informs whether modeling

dynamic changes during CBT-I is clinically relevant, above and beyond simple pre-post intervention analyses.

We evaluated the model fit indices for the latent growth curve models which included treatment response as a predictor (Table S7). The models for all sleep variables, except for sleep latency, had acceptable model fit indices.

Table 6 (shown on the next page) shows the dynamic changes during CBT-I to predict treatment outcome. Baseline (i.e., intercept) sleep characteristics and baseline pre-sleep arousal were not substantially associated with treatment outcome (βs $\leq$ |0.25|). This suggests that participant's pre-intervention values on these measures did not predict whether or not they responded to CBT-I.

The rate (slope) and shape (quadratic) of change for some sleep characteristics *were* significantly associated with treatment outcome, indicated by the ISI change score, with higher values indicating a greater reduction in insomnia symptoms. Faster initial improvements in sleep efficiency ($\beta = 0.34$) and sleep quality ($\beta = 0.54$), as well as a smaller initial decrease in alertness ($\beta = 0.45$), were associated with a larger improvement in insomnia symptoms. Maintaining higher levels of sleep quality ($\beta = -0.50$) and sleep efficiency ($\beta = -0.33$), and quicker subsequent improvements in alertness ($\beta = -0.39$) during the intervention (i.e., a smaller quadratic term) was associated with a larger improvement in insomnia symptoms. Rate and shape of change in WASO, sleep duration, and pre-sleep arousal were not associated with treatment outcomes (βs $\leq$ |0.22|). Thus, taken together, larger linear improvements and less leveling off of changes were associated with better treatment outcomes.

Table 6 Dynamic changes during CBT-I to predict treatment outcome

	WASO	Sleep efficiency	Sleep duration	Alertness	Quality	Pre-sleep arousal
Regression Slopes, Beta(SE)						
Intercept → ISI	-0.15(0.11)	0.17(0.11)	0.25(0.1)	0.17(0.11)	0.13(0.11)	0.09(0.11)
Slope → ISI	-0.15(0.13)	0.34(0.11)*	0.12(0.13)	0.45(0.11)*	0.54(0.09)*	-0.22(0.13)
Quadratic → ISI	0.15(0.14)	-0.33(0.13)*	-0.12(0.14)	-0.39(0.15)*	-0.50 (0.11)*	0.1(0.15)
Latent Intercepts						
Intercept, B(SE)	86.90(12.27)*	74.64(2.96)*	351.99(15.89)*	51.87(5.10)*	45.32(4.16)*	28.97(3.77)*
Linear Slope, Beta(SE)	-1.0(0.36)*	0.52(0.34)	-1.11(0.35)*	-1.19(0.29)*	-0.11(0.28)	-0.16(0.35)
Quadratic Slope, Beta(SE)	0.78(0.38)	-0.24(0.35)	1.18(0.39)*	1.35(0.39)*	-0.34(0.3)	0.24(0.39)
Latent Covariances, Beta(SE)						
Intercept w/Linear Slope	-0.78(0.06)*	-0.53(0.1)*	-0.46(0.11)*	-0.03(0.16)	-0.4(0.12)*	-0.32(0.13)
Intercept w/Quadratic Slope	0.69(0.09)*	0.40(0.13)*	0.37(0.13)*	-0.07(0.19)	0.30(0.14)	0.35(0.15)
Linear Slope w/Quadratic Slope	-0.96(0.01)*	-0.96(0.02)*	-0.98(0.01)*	-0.93(0.03)*	-0.91(0.03)*	-0.93(0.03)*

Notes. WASO, wake after sleep onset; ISI, Insomnia Severity Index Change Score; B, unstandardized coefficients; SE, standard error; Beta, standardized coefficient; * p <.01

4.0 Discussion

This study is the first to evaluate linear and non-linear changes in sleep characteristics and pre-sleep arousal *during* CBT-I to predict treatment outcomes. We report that rates of change for some sleep characteristics (sleep efficiency, sleep quality, and alertness) were associated with greater treatment response, whereas pre-intervention sleep characteristics were not predictive of treatment outcome. This study suggests that evaluating dynamic changes during CBT-I may be more clinically informative than pre-post-intervention analyses alone. Evaluating these changes is feasible during the course of clinical care, as daily sleep diaries are a standard of CBT-I treatment. Additionally, our research methods more closely matches routine clinical practice, because clinicians review the sleep diaries at the beginning of each CBT-I session to monitor progress and modify treatment plans.

4.1 Several changes in sleep characteristics during CBT-I predicted treatment response

CBT-I is an excellent example of measurement-based care – that is, the clinician uses data collected *during* the intervention to modify treatment (Fortney et al., 2017). However, most of the research on CBT-I interventions have not captured these dynamic processes. Only one previous study has evaluated changes in sleep characteristics during CBT-I to predict treatment outcomes. Morin and colleagues (2014) divided their participants into tertiles (worst change, average change, and best change) for five sleep characteristics from baseline to the first week of the intervention. They reported that patient group membership did not predict treatment response

for any sleep characteristic (Morin et al., 2014). In contrast to their findings, we demonstrated that several sleep characteristics were significant predictors of treatment response. This difference in results is likely because of our greater statistical power. Our methods allowed for linear and non-linear changes for all eight weeks of the intervention and for the full sample of participants.

In the current study, we observed that: (1) improvements in sleep efficiency were strongly predictive of treatment success; (2) individuals who experienced larger initial decrements in alertness had worse treatment response; and (3) changes in sleep duration were not predictive of treatment response. This pattern of results is important to examine together, as they are all associated with the recommendation that patients restrict their time in bed (Perlis et al., 2021). Time in bed restriction increases individuals' homeostatic sleep pressure (Maurer et al., 2022), which helps patients to fall asleep and stay asleep more easily. However, time in bed restriction can also reduce total sleep time and alertness, which may be undesirable for patients and could increase fall risk among older adults (Hughes & Martin, 2021; Kyle et al., 2014). Therefore, determining optimal time in bed restriction recommendations is a balancing act of these competing concerns – maximizing sleep efficiency improvements while minimizing risk of alertness decrements.

Our results suggest that clinicians should attempt a strong time in bed restriction (e.g., sleep duration plus 30 minutes) to obtain the optimal improvement in sleep efficiency, which is the most potent predictor of treatment success in our study. These results also suggest that clinicians should discuss with the patient that this reduction of time in bed may lead to decrements in alertness, which is associated with a poor treatment outcome. To mitigate this issue, the clinician might recommend a strategically timed nap (i.e., in the afternoon, for

approximately 45 minutes). This recommendation does not have a large compromising effect on the accumulation of sleep drive and does not induce any circadian phase shifts (Evans & Hasler, 2021). Finally, our results suggest that the patient – and the clinician – should not be overly concerned with changes to sleep duration, as this does not predict treatment success in our study.

Future studies should test varying time in bed restriction methods to optimally balance sleep efficiency and alertness changes during CBT-I. Time in bed restriction may be the most effective component of CBT-I, though no dismantling studies have compared time in bed restriction therapy alone to multi-component CBT-I to test this hypothesis (Perlis et al., 2021). Among older adults, a more modest restriction of time in bed has been suggested, such as reducing the sleep window by 30-minutes per week until sleep efficiency improves (Hughes & Martin, 2021). This contrasts with concerns that a conservative approach to time in bed restriction ("under-titrating") may lead to both poor treatment response as well as patient dissatisfaction when they do not experience rapid improvements (Perlis, 2017). These future studies would improve precision of clinical decision-making about optimal time in bed restriction recommendations.

4.2 Pre-sleep arousal reductions did not predict treatment outcome

We evaluated changes in pre-sleep arousal *during* CBT-I to improve our understanding of how one treatment marker unfolds over time. We reported that CBT-I was associated with an initial medium-sized reduction in pre-sleep arousal, consistent with previous pre-post-intervention studies (Vincent & Lewycky, 2009; Wu et al., 2006). In multi-component CBT-I, pre-sleep arousal may be affected by cognitive therapy addressing maladaptive beliefs about

sleep, as well as stimulus control, which reduces the association of the bed being a place of worry.

In our bi-directional models, increases in alertness (upon awakening) were associated with subsequent decreases in pre-sleep arousal (occurring at bedtime). This may be because better alertness is associated with better mood (Johnson et al., 1990) and better cognitive performance (Rogers et al., 2003), thereby reducing pre-sleep arousal. This finding, in conjunction with our finding that declines in alertness are predictive of treatment outcome, underscores how critical it is for clinicians to assess and monitor alertness during CBT-I. Alertness is assessed in some (e.g., Consensus Sleep Diary, Carney et al., 2011), but not all (e.g., sleep raster plot) sleep diaries used during CBT-I.

Though we hypothesized that there would be a bidirectional relationship between sleep and pre-sleep arousal, pre-sleep arousal changes were not associated with subsequent sleep characteristic changes. This finding is consistent with a previous systematic review which indicated that while sleep changes reliably influence next-day negative affect, there was inconsistent and weaker evidence that daytime affect changes influence that night's sleep (Kahn et al., 2013). Additionally, there are individual differences in the degree to which stressful events precipitate sleep disturbances (Drake et al., 2004; Kalmbach et al., 2018). In the context of CBT-I, our results suggest that pre-sleep arousal reductions might not occur until later in treatment, when the sleep window has been expanded and alertness increases. Additionally, pre-sleep arousal reductions may occur later in treatment because cognitive therapy techniques are typically introduced in later sessions. Future work might examine the precise timing of pre-sleep arousal improvement (e.g., Batterham et al., 2017), or conduct a dismantling study to determine

if behavioral or cognitive therapies have more of an impact (e.g., Harvey et al., 2014) to better understand the dynamics of pre-sleep arousal changes during treatment.

While rates of change of several sleep characteristics predicted treatment outcomes in our study, the rate of change in pre-sleep arousal did not. Previous studies have identified that pre-sleep arousal (change from pre-post intervention) is a mediator of the association between CBT-I and treatment outcomes (Parsons et al., 2021; Schwartz & Carney, 2012). This discrepancy with our results may be due to our different methods of modeling change in pre-sleep arousal, as well as the fact that we did not have a control group with which to compare our changes. Additionally, in our study the level of pre-sleep arousal was somewhat low at baseline, with an average participant rating of 31.5 (with 100 being the highest arousal rating). If our results were to be replicated, it would suggest that pre-sleep arousal is not a target for measurement-based care during CBT-I, since its change (or lack thereof) during the intervention is not predictive of treatment outcome.

Clinicians and researchers may also find that pre-sleep arousal is an important variable to target during treatment for some, but not all, patients. Two previous studies of dynamic change during CBT-I identified that there were distinct subgroups of depressive symptom trajectories. One of these studies identified two groups, one that improved, and one that was a mixture of individuals with no change or a decline in symptoms (Batterham et al., 2017). The other study identified three groups of depressive symptom change: initial, partial, and optimal responders (Bei et al., 2018). It is possible that the decelerating rate of change (i.e., leveling off) in pre-sleep arousal towards the end of treatment in our study may be representing initial or partial responders, and there may be other individuals who have a continual reduction in pre-sleep arousal. Though the average in our sample for pre-sleep arousal was 31.5, there was a large

range of scores of 3-72, suggesting that this is a salient variable for some of our participants, but not all. The implementation of latent class analysis to identify subgroups on the basis of change in pre-sleep arousal may further improve measurement-based care and precision medicine efforts in CBT-I.

4.3 Limitations

Despite this research study more closely aligning to routine clinical care than previous work, there are several limitations. First, this study was not originally designed for our study aims. Our data analysis did not include a control group, and so we could not compare CBT-I to a "treatment as usual" group. We could also not adequately model the rate of change of sleep latency, likely because sleep latency has the greatest instability in both individuals with insomnia and controls and required more participants (Wohlgemuth et al., 1999). More participants would have also allowed us to analyze the data at the weekly-session level, rather than bi-weekly, which would more closely match clinical care. Given power limitations, our change in insomnia symptoms was measured as a continuous change score, rather than a binary outcome (e.g., responder, non-responder).

Second, the generalizability of this study was limited by several of its demographic features. Our study was conducted among older adults, and therefore is not generalizable to younger samples, who have been shown to have even larger responses to CBT-I (Irwin et al., 2006; van Straten et al., 2018). The intervention was not modified for older adults' needs in advance (Hughes & Martin, 2021), though specific recommendations were tailored to the individuals' needs and presenting concerns. Only 30% of our participants were male, and an

even smaller proportion were from racial/ethnic minority groups (nine African Americans and two Asian Americans in total, no Hispanic individuals). It is possible that this sample is representative of those who present to clinics with insomnia, as females and Caucasians are more likely to report symptoms of insomnia than their male and non-White counterparts, respectively (Mallampalli & Carter, 2014; Ruiter et al., 2010). However, the effects of confounding by cultural differences in these groups' beliefs, values, and attitudes towards sleep have not yet been fully described (Shaw et al., 2012). Therefore, our results cannot be generalized to all demographics of a culturally and ethnically diverse general population.

Third, though we have conceptualized changes in sleep characteristics as a marker of treatment, one might argue that they are measures of adherence. In the current study, we had a measure of therapist-reported patient engagement and understanding of the session materials (both on a scale of 0-3, with 3 indicating highest levels). Because more than 90% of patient sessions were rated at a 3 (maximum), and the remaining 10% were a 2 (moderate levels of engagement or understanding), there was insufficient heterogeneity to assess whether these therapist-reported measures were moderators of observed changes. The measures of adherence that are typically extracted from daily sleep diaries during CBT-I are bedtime variability, risetime variability, getting out of bed if one is not sleeping, and time in bed (Dong et al., 2018). These measures are all more proximal indicators of adherence to behavioral recommendations during CBT-I (reduce time in bed, maintain a consistent wake-up time) than the sleep characteristics we assessed. Additionally, the sleep characteristics in the current study have been previously identified as (secondary) treatment outcomes in CBT-I clinical guidelines (Edinger et al., 2021). We maintain that the sleep characteristics that we examined were treatment markers and not adherence measures because clinicians cannot instruct patients to fall asleep more

quickly or become more alert. Rather, the clinician provides cognitive and behavioral recommendations, and if the patient adheres to these, they may subsequently experience improvements in sleep latency and alertness.

4.4 Conclusion

Our study demonstrated linear and non-linear changes in sleep characteristics during CBT-I that were consistent with our current understanding of the etiology and cognitive-behavioral treatment of insomnia. We demonstrated that these changes in sleep characteristics during CBT-I predicted treatment outcomes, whereas pre-intervention data did not predict treatment success. Thus, CBT-I research may benefit from assessing diary data collected during the intervention. Such a change also aligns more closely with clinical care, which would make this research even more translational.

We reported that improvements in sleep efficiency predicted treatment success, while decrements in alertness portended poor outcomes. These changes, if replicated in diverse samples, call for monitoring these characteristics and modifying treatment recommendations accordingly. Clinicians have long recognized that changes in sleep efficiency are a key marker to success during CBT-I, and so this finding is consistent with that focus. However, alertness is not consistently monitored during the course of CBT-I, and our results suggest that doing so may be critical. Including a measure of alertness during CBT-I would also provide another indicator to aid in clinical decision-making about whether the time in bed restriction should be modified.

This study aimed to improve the precision of treatment recommendations during CBT-I through the use of daily diary data. It is critical to continue to refine the precision of CBT-I, because pre-post intervention analyses suggest that it only causes insomnia remission among 50% of patients (Edinger et al., 2021). Future research may increase the success of CBT-I to 100% of patients by analyzing diary data collected during the course of treatment (as we did here), evaluating which subgroups of patients are particularly responsive or non-responsive to treatment, experimentation with varying degrees of time in bed restriction in samples of older and younger adults, and dismantling studies to best understand the impact of each component of this multi-component treatment.

CBT-I has been recognized since 2016 by internal medicine physicians as the frontline treatment for individuals with insomnia (Qaseem et al., 2016). It is our duty as psychologists to continue to improve CBT-I to reduce insomnia symptoms for the millions of Americans currently suffering from sleepless nights.

5.0 Supplemental Tables

Table S1 Between-person correlations of diary variables

	Sleep latency	WASO	Sleep efficiency	Sleep duration	Alertness	Sleep quality
Baseline						
WASO	0.25					
Sleep efficiency	-0.61*	-0.88*				
Sleep duration	-0.37	-0.50*	0.71*			
Alertness	-0.20	-0.05	0.10	0.01		
Sleep quality	-0.48*	-0.57*	0.65*	0.39	0.35	
Pre-sleep arousal	-0.01	0.04	0.00	0.04	-0.02	0.00
Intervention						
WASO	0.31					
Sleep efficiency	-0.63*	-0.92*				
Sleep duration	-0.46*	-0.68*	0.79*			
Alertness	-0.23	-0.16	0.22	0.21		
Sleep quality	-0.40*	-0.65*	0.70*	0.67*	0.53*	
Pre-sleep arousal	0.18	-0.13	0.02	-0.02	-0.18	-0.24
Post-Intervention						
WASO	0.70*					
Sleep efficiency	-0.79*	-0.87*				
Sleep duration	-0.38	-0.39	0.60*			
Alertness	-0.15	-0.11	0.22	0.13		
Sleep quality	-0.47	-0.56*	0.59*	0.35	0.59*	
Pre-sleep arousal	0.32	0.28	-0.39	-0.18	-0.36*	-0.63*

Notes. WASO, wake after sleep onset; * indicates $p < .0001$

Table S2 Repeated measures correlations of diary variables

	Sleep latency	WASO	Sleep efficiency	Sleep duration	Alertness	Sleep quality
Baseline						
WASO	0.02					
Sleep efficiency	-0.54*	-0.79*				
Sleep duration	-0.37*	-0.55*	0.79*			
Alertness	-0.06	-0.13*	0.19*	0.19*		
Sleep quality	-0.32*	-0.41*	0.53*	0.40*	0.30*	
Pre-sleep arousal	0.30*	0.16*	-0.30*	-0.24*	-0.12*	-0.33*
Intervention						
WASO	0.13*					
Sleep efficiency	-0.62*	-0.80*				
Sleep duration	-0.35*	-0.34*	0.61*			
Alertness	-0.09*	-0.17*	0.20*	0.20*		
Sleep quality	-0.35*	-0.51*	0.60*	0.33*	0.32*	
Pre-sleep arousal	0.29*	0.23*	-0.33*	-0.20*	-0.12*	-0.36*
Post-Intervention						
WASO	0.40*					
Sleep efficiency	-0.80*	-0.85*				
Sleep duration	-0.68*	-0.67*	0.85*			
Alertness	-0.28*	-0.26*	0.26*	0.28*		
Sleep quality	-0.35*	-0.59*	0.61*	0.50*	0.43*	
Pre-sleep arousal	0.35*	0.17*	-0.31*	-0.25*	-0.19*	-0.34*

Notes. WASO, wake after sleep onset; * indicates $p < .0001$

Table S3 Percentage of within-person variance

	Sleep latency	WASO	Sleep efficiency	Sleep duration	Alertness	Sleep quality	Pre-sleep arousal
Baseline	69.0	59.7	59.7	64.5	47.6	63.9	61.1
Intervention	61.4	58.0	52.9	57.1	38.0	54.7	53.5
Post-Intervention	64.3	62.9	74.5	72.3	26.3	45.1	41.4

Notes. WASO, wake after sleep onset. These percentages are calculated using the intraclass correlation coefficient

Table S4 Fit indices of latent growth curve models

		χ^2	AIC	BIC	CFI	TLI	RMSEA	SRMR
Sleep latency	Linear	267.00(21)	4493.39	4508.65	0.62	0.73	0.35	0.21
	Quadratic	186.87(17)	4421.26	4446.69	0.74	0.77	0.33	0.19
WASO	Linear	211.83(21)	5004.38	5019.64	0.59	0.70	0.31	0.19
	Quadratic	67.29(17)	4867.83	4893.26	0.89	0.90	0.18	0.11
Sleep efficiency	Linear	175.41(21)	3468.57	3483.83	0.76	0.83	0.28	0.11
	Quadratic	44.36(17)	3345.53	3370.97	0.96[+]	0.96[+]	0.13	0.04[+]
Sleep duration	Linear	199.17(21)	5361.3	5376.56	0.68	0.77	0.30	0.13
	Quadratic	86.44(17)	5256.57	5282	0.87	0.89	0.21	0.08[+]
Alertness	Linear	84.00(21)	4201.68	4216.87	0.90	0.93[+]	0.18	0.1
	Quadratic	43.13(17)	4168.81	4194.13	0.96[+]	0.97[+]	0.13	0.06[+]
Sleep quality	Linear	119.33(21)	4045	4060.2	0.82	0.87	0.22	0.1
	Quadratic	24.80(17)	3958.47	3983.8	0.99[+]	0.99[+]	0.07	0.06[+]
Pre-sleep arousal	Linear	79.01(21)	3954.04	3969.23	0.88	0.91	0.17	0.07
	Quadratic	41.57(17)	3924.6	3949.92	0.95[+]	0.96[+]	0.12	0.04[+]

Notes WASO, wake after sleep onset; χ2 , chi-squared statistic; AIC, Akaike's Information Criteria; BIC, Bayesian Information Criteria; CFI, Comparative Fit Index; TLI, Tucker Lewis Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual; [+] indicates acceptable fit

Table S5 Fit indices for multivariate latent growth curve models

	χ^2	AIC	BIC	CFI	TLI	RMSEA	SRMR
WASO + Sleep duration	218.07(61)	10070.96	10144.71	0.86	0.85	0.17	0.09
WASO + Sleep quality	157.75(61)	8750.62	8824.38	0.91	0.91	0.13	0.08^+
Sleep efficiency + Alertness	105.85(61)	7512.93	7586.69	0.97^+	0.96^+	0.09	0.04^+
Sleep duration + Alertness	146.18(61)	9426.45	9500.2	0.93	0.92	0.12	0.06^+
Sleep duration + Sleep quality	163.60(61)	9173.95	9247.7	0.91	0.90	0.13	0.06^+

Notes. WASO, wake after sleep onset; χ2 , chi-squared statistic; AIC, Akaike's Information Criteria; BIC, Bayesian Information Criteria; CFI, Comparative Fit Index; TLI, Tucker Lewis Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual. $^+$ indicates acceptable model fit

Table S6 Fit indices for random-intercept cross-lagged panel models

	χ^2	AIC	BIC	CFI	TLI	RMSEA	SRMR
Sleep Latency	214.40(53)	8296.86	8390.96	0.86	0.83	0.18	0.18
WASO	103.60(53)	8776.15	8870.25	0.95^+	0.93	0.1	0.11
Sleep Efficiency	87.99(53)	7255.19	7349.29	0.97^+	0.96^+	0.08	0.14
Sleep Duration	123.00(53)	9153.77	9247.87	0.93	0.92	0.12	0.18
Quality	100.15(53)***	7860.47	7954.17	0.96^+	0.95^+	0.1	0.18
Alertness	88.37(53)**	8084.86	8178.57	0.97^+	0.96^+	0.08	0.11

Notes. WASO, wake after sleep onset; χ2 , chi-squared statistic; AIC, Akaike's Information Criteria; BIC, Bayesian Information Criteria; CFI, Comparative Fit Index; TLI, Tucker Lewis Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual. $^+$ indicates acceptable model fit

Table S7 Fit indices for models predicting treatment outcome

	χ^2	AIC	BIC	CFI	TLI	RMSEA	SRMR
SOL	187.71(20)	4337.42	4370.2	0.74	0.73	0.3	0.17
WASO	75.44(20)	4761.9	4794.68	0.88	0.87	0.17	0.08[+]
SE	47.63(20)	3264.03	3296.81	0.96[+]	0.95[+]	0.12	0.03[+]
TST	94.50(20)	5152.94	5185.73	0.87	0.86	0.2	0.07[+]
Alertness	61.88(20)	4064.8	4097.59	0.94	0.94	0.15	0.05[+]
Quality	38.68(20)	3834.22	3867	0.97[+]	0.97[+]	0.1	0.05[+]
PSA	56.83(20)	3853	3885.78	0.93	0.92	0.14	0.04[+]

Notes. SOL, sleep onset latency; WASO, wake after sleep onset; SE, sleep efficiency; TST, total sleep time; PSA, pre-sleep arousal; χ2 , chi-squared statistic; AIC, Akaike's Information Criteria; BIC, Bayesian Information Criteria; CFI, Comparative Fit Index; TLI, Tucker Lewis Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual; [+] indicates acceptable fit

6.0 Bibliography

American Psychiatric Association. (2013a). Insomnia disorder. In *Diagnostic and statistical manual of mental disorders* (5th ed.). https://10.1176/appi.books.9780890425596.744053

American Psychiatric Association. (2013b). Anxiety disorders. In *Diagnostic and statistical manual of mental disorders* (5th ed.). https://10.1176/appi.books.9780890425596.744053

Ancoli-Israel, S., & Martin, J. L. (2006). Insomnia and daytime napping in older adults. *Journal of Clinical Sleep Medicine*, *2*(3), 333–342. https://doi.org/10.5664/jcsm.26597

Auguie, B. (2017). gridExtra: Miscellaneous functions for "grid" graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, *8*, 456. https://doi.org/10.3389/fpsyg.2017.00456

Bakdash, J. Z., & Marusich, L. R. (2021). rmcorr: Repeated measures correlation. R package version 0.4.3. https://CRAN.R-project.org/package=rmcorr

Bathgate, C. J., Edinger, J. D., & Krystal, A. D. (2017). Insomnia patients with objective short sleep duration have a blunted response to cognitive behavioral therapy for insomnia. *Sleep*, *40*(1), zsw012. http://dx.doi.org/10.1093/sleep/zsw012

Batterham, P. J., Christensen, H., Mackinnon, A. J., Gosling, J. A., Thorndike, F. P., Ritterband, L. M., Glozier, N., & Griffiths, K. M. (2017). Trajectories of change and long-term outcomes in a randomised controlled trial of internet-based insomnia treatment to prevent depression. *BJPsych Open*, *3*(5), 228–235. https://doi.org/10.1192/bjpo.bp.117.005231

Bastien, C. H., Vallières, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Medicine*, *2*(4), 297–307.

https://doi.org/10.1080/15402002.2011.606766

Bei, B., Asarnow, L. D., Krystal, A., Edinger, J. D., Buysse, D. J., & Manber, R. (2018). Treating insomnia in depression: insomnia related factors predict long-term depression trajectories. *Journal of Consulting and Clinical Psychology*, *86*(3), 282–293. https://doi.org/10.1037/ccp0000282

Benz, F., Knoop, T., Ballesio, A., Bacaro, V., Johann, A. F., Rücker, G., Feige, B., Riemann, D., & Baglioni, C. (2020). The efficacy of cognitive and behavior therapies for insomnia on daytime symptoms: A systematic review and network meta-analysis. *Clinical Psychology Review*, *80*. https://doi.org/10.1016/j.cpr.2020.101873

Bertisch, S. M., & Buysse, D. J. (2021). Pharmacologic treatment I: Therapeutic approaches and implementation. In M. Krieger, T. Roth, C. Goldstein, W. Dement (Eds.), *Principles and practice of sleep medicine* (7th ed., Vol. 1, pp. 902-910). Elsevier.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Press.

Borbély, A. A., Daan, S., Wirz-Justice, A., & Deboer, T. (2016). The two-process model of sleep regulation: A reappraisal. *Journal of Sleep Research*, *25*(2), 131–143. https://doi.org/10.1111/jsr.12371

Byrne, P. R. (2004). Anxiety. In W. E. Craighead & C. B. Nemeroff (Eds.), *The concise Corsini encyclopedia of psychology and behavioral science*. John Wiley & Sons.

Carney, C. E., & Danforth, M. (2021). Behavioral treatment I: Therapeutic approaches and implementation. In M. Krieger, T. Roth, C. Goldstein, W. Dement (Eds.), *Principles and Practice of Sleep Medicine* (7th ed., Vol. 1, pp. 883-890). Elsevier.

Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., ... & PROMIS Cooperative

Group. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*, *63*(11), 1179-1194. https://doi.org/10.1016/j.jclinepi.2010.04.011

Chakravorty, S., Morales, K. H., Arnedt, J. T., Perlis, M. L., Oslin, D. W., Findley, J. C., & Kranzler, H. R. (2019). Cognitive behavioral therapy for insomnia in alcohol-dependent veterans: a randomized, controlled pilot study. *Alcoholism: Clinical and Experimental Research*, *43*(6), 1244-1253. https://doi.org/10.1111/acer.14030.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). Perceived stress scale (PSS). *Journal of Health and Social Behavior*, *24*, 385-396.

Curran, P. J., Howard, A. L., Bainter, S., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, *82*(5), 879–894. https://doi.org/10.1037/a0035297

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121–136. https://doi.org/10.1080/15248371003699969

de Almondes, K. M., Costa, M. V., Malloy-Diniz, L. F., & Diniz, B. S. (2016). Insomnia and risk of dementia in older adults: Systematic review and meta-analysis. *Journal of Psychiatric Research*, *77*, 109–115. https://doi.org/10.1016/j.jpsychires.2016.02.021

Dong, L., Soehner, A. M., Bélanger, L., Morin, C. M., & Harvey, A. G. (2018). Treatment agreement, adherence, and outcome in cognitive behavioral treatments for insomnia. *Journal of Consulting and Clinical Psychology, 86*(3), 294. https://doi.org/

10.1037/ccp0000269

Drake, C., Richardson, G., Roehrs, T., Scofield, H., & Roth, T. (2004). Vulnerability to stress-related sleep disturbance and hyperarousal. *Sleep*, *27*(2), 285-291.

Edinger, J. D., Arnedt, J. T., Bertisch, S. M., Carney, C. E., Harrington, J. J., Lichstein, K. L., Sateia, M. J., Troxel, W. M., Zhou, E. S., Kzmi, U., Heald, J. L., & Martin, J. L. (2021). Behavioral and psychological treatments for chronic insomnia disorder in adults: an American Academy of Sleep Medicine systematic review, meta-analysis and GRADE assessment. *Journal of Clinical Sleep Medicine*, 17(2), 255-262. https://doi.org/10.5664/jcsm.8986

Edinger, J. D., & Carney, C. E. (2015). *Treatments that work. Overcoming insomnia: A cognitive-behavioral therapy approach, therapist guide (2nd ed.).* Oxford University Press.

Espie, C. A., Inglis, S. J., & Harvey, L. (2001). Predicting clinically significant response to cognitive behavior therapy for chronic insomnia in general medical practice: Analyses of outcome data at 12 months posttreatment. *Journal of Consulting and Clinical Psychology*, *69*(1), 58–66. https://doi.org/10.1037/0022-006X.69.1.58

Evans, M. A., & Hasler, B. P. (2021). CBT-I for patients with phase disorders or insomnia with circadian misalignment. In S. Nowakowski, S. Garland, M. Grandner, & L. Cuddihy (Eds.), *Adapting cognitive behavioral therapy for insomnia* (pp. 63-95). Academic Press.

Flourney, J. Generate lavaan syntax for RI-CLPM. Retreived on November 10, 2021 from: jflournoy.github.io/riclpmr/

Fortier-Brochu, E., Beaulieu, Ivers, H., & Morin, C. M. (2012). Insomnia and daytime cognitive performance: A meta-analysis. *Sleep Medicine Reviews*, *16*(1), 83–94.

https://doi.org/10.1016/j.smrv.2011.03.008

Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). A tipping point for measurement-based care. *Psychiatric Services*, *68*(2), 179–188. https://doi.org/10.1176/appi.ps.201500439

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

Gooneratne, N. S., Bellamy, S. L., Pack, F., Staley, B., Schutte-Rodin, S., Dinges, D. F., & Pack, A. I. (2011). Case-control study of subjective and objective differences in sleep patterns in older adults with insomnia symptoms. *Journal of Sleep Research*, *20*(3), 434–444. https://doi.org/10.1111/j.1365-2869.2010.00889.x

Grandner, M. A., Patel, N. P., Gehrman, P. R., Perlis, M. L., & Pack, A. I. (2010). Problems associated with short sleep: Bridging the gap between laboratory and epidemiological studies. *Sleep Medicine Reviews*, *14*(4), 239–247. https://doi.org/10.1016/j.smrv.2009.08.001

Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software, 40*(3), 1-25. https://www.jstatsoft.org/v40/i03/.

Hall, M. H., Fernandez-Mendoza, J., Kline, C. E., Evans, M. A., & Vgontzas, A. (2021). Insomnia with short sleep duration and multidimensional sleep health. In M. Krieger, T. Roth, C. Goldstein, W. Dement (Eds.), *Principles and practice of sleep medicine* (7th ed., Vol. 1, pp. 867-882). Elsevier.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Harrell, F. E. (2021). Hmisc: Harrell Miscellaneous. R package version 4.5-0. https://CRAN.R-project.org/package=Hmisc

Harvey, A. G. (2005). A cognitive theory and therapy for chronic insomnia. *Journal of Cognitive Psychotherapy*, *19*(1), 41-59. https://doi.org/10.1891/jcop.19.1.41.66332

Harvey, A. G., Bélanger, L., Talbot, L., Eidelman, P., Beaulieu-Bonneau, S., Fortier-Brochu, É., Ivers, H., Lamy, M., Hein, K., Soehner, A. M., Mérette, C., & Morin, C. M. (2014). Comparative efficacy of behavior therapy, cognitive therapy, and cognitive behavior therapy for chronic insomnia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *82*(4), 670. https://doi.org/ 10.1037/a0036606

Hertenstein, E., Feige, B., Gmeiner, T., Kienzler, C., Spiegelhalder, K., Johann, A., Jansson-Fröjmark, M., Palagini, L., Rücker, G., Riemann, D., & Baglioni, C. (2019). Insomnia as a predictor of mental disorders: A systematic review and meta-analysis. *Sleep Medicine Reviews*, *43*, 96–105. https://doi.org/10.1016/j.smrv.2018.10.006

Hlavac, Marek (2018). Stargazer: Well-formatted regression and summary statistics tables. R package version 5.2.1. https://CRAN.R-project.org/package=stargazer

Hoberg, K., Badorf, F., & Lapp, L. (2017). The inverse hockey stick effect: An empirical investigation of the fiscal calendar's impact on firm inventories. *International Journal of Production Research*, *55*(16), 4601-4624. https://doi.org/10.1080/00207543.2016.1269969

Hope, R. M. (2013). Rmisc: Rmisc: Ryan Miscellaneous. R package version 1.5. https://CRAN.R-project.org/package=Rmisc

Hu, L. T., & Bentler, P. M. (2009). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A*

Multidisciplinary Journal, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hughes, J. M. & Martin, J. L. (2021). CBT-I for older adults. In S. Nowakowski, S. Garland, M. Grandner, & L. Cuddihy (Eds.), *Adapting cognitive behavioral therapy for insomnia* (pp. 347-368)*.* Academic Press.

Irwin, M. R., Cole, J. C., & Nicassio, P. M. (2006). Comparative meta-analysis of behavioral interventions for insomnia and their efficacy in middle-aged adults and in older adults 55+ years of age. *Health Psychology*, *25*(1), 3–14. https://doi.org/10.1037/0278-6133.25.1.3

Itani, O., Jike, M., Watanabe, N., & Kaneita, Y. (2017). Short sleep duration and health outcomes: A systematic review, meta-analysis, and meta-regression. *Sleep Medicine*, *32*, 246–256. https://doi.org/10.1016/j.sleep.2016.08.006

Jansson-Fröjmark, M., & Norell-Clarke, A. (2012). Psychometric properties of the Pre-Sleep Arousal Scale in a large community sample. *Journal of Psychosomatic Research*, *72*(2), 103–110. https://doi.org/10.1016/j.jpsychores.2011.10.005

Johns, M. W.(1991). A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *Sleep,* 14, 540–545.

Johnson, L. C., Spinweber, C. L., Gomez, S. A., & Matteson, L. T. (1990). Daytime sleepiness, performance, mood, nocturnal sleep: the effect of benzodiazepine and caffeine on their relationship. *Sleep*, *13*(2), 121-135.

Johnson, P. & Kite, B. (2020). semTable: Structural Equation Modeling Tables. R package version 1.8. https://CRAN.R-project.org/package=semTable

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). semTools: Useful tools for structural equation modeling. R package version 0.5-5. Retrieved from

https://CRAN.R-project.org/package=semTools

Kahn, M., Sheppes, G., & Sadeh, A. (2013). Sleep and emotions: Bidirectional links and underlying mechanisms. *International Journal of Psychophysiology*, *89*(2), 218–228. https://doi.org/10.1016/j.ijpsycho.2013.05.010

Kalmbach, D. A., Anderson, J. R., & Drake, C. L. (2018). The impact of stress on sleep: Pathogenic sleep reactivity as a vulnerability to insomnia and circadian disorders. *Journal of Sleep Research*, *27*(6). https://doi.org/10.1111/jsr.12710

Kay, D. B., Buysse, D. J., Germain, A., Hall, M., & Monk, T. H. (2015). Subjective-objective sleep discrepancy among older adults: Associations with insomnia diagnosis and insomnia treatment. *Journal of Sleep Research*, *24*(1), 32–39. https://doi.org/10.1111/jsr.12220

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kyle, S. D., Miller, C. B., Rogers, Z., Siriwardena, A. N., Macmahon, K. M., & Espie, C. A. (2014). Sleep restriction therapy for insomnia is associated with reduced objective total sleep time, increased daytime somnolence, and objectively impaired vigilance: Implications for the clinical management of insomnia disorder. *Sleep*, *37*(2), 229–237. https://doi.org/10.5665/sleep.3386

Lefcheck, J. (2021). Multigroup analysis. In *Structural Equation Modeling in R for Ecology and Evolution*. https://jslefche.github.io/sem_book/index.html

Mallampalli, M. P., & Carter, C. L. (2014). Exploring sex and gender differences in sleep health: A society for women's health research report. *Journal of Women's Health*, *23*(7), 553–

562. https://doi.org/10.1089/jwh.2014.4816

Maurer, L. F., Espie, C. A., Omlin, X., Emsley, R., & Kyle, S. D. (2022). The effect of sleep restriction therapy for insomnia on sleep pressure and arousal: A randomized controlled mechanistic trial. *Sleep*, *45*(1), 1–12. https://doi.org/10.1093/sleep/zsab223

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, *28*(6), 487-495.

Mitchell, L. J., Bisdounis, L., Ballesio, A., Omlin, X., & Kyle, S. D. (2019). The impact of cognitive behavioural therapy for insomnia on objective sleep parameters: A meta-analysis and systematic review. *Sleep Medicine Reviews*, *47*(0), 90–102. https://doi.org/10.1016/j.smrv.2019.06.002

Monk, T. H., Reynolds III, C. F., Kupfer, D. J., Buysse, D. J., Coble, P. A., Hayes, A. J., Machen, M. A., Petrie, S. R., & Ritenour, A. M. (1994). The Pittsburgh Sleep Diary. *Journal of Sleep Research*, *3*(2), 111-120.

Morin, C. M., Beaulieu-Bonneau, S., Ivers, H., Vallières, A., Guay, B., Savard, J., & Mérette, C. (2014). Speed and trajectory of changes of insomnia symptoms during acute treatment with cognitive-behavioral therapy, singly and combined with medication. *Sleep Medicine*, *15*(6), 701–707. https://doi.org/10.1016/j.sleep.2014.02.004

Morin, C. M., Vallières, A., & Ivers, H. (2007). Dysfunctional beliefs and attitudes about sleep (DBAS): Validation of a brief version (DBAS-16). *Sleep*, *30*(11), 1547–1554. https://doi.org/10.1093/sleep/30.11.1547

Mulder, J. D., & Hamaker, E. L. (2021). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(4), 638-648.

https://doi.org/10.1080/10705511.2020.1784738

Mund, M., & Nestler, S. (2019). Beyond the cross-lagged panel model: Next-generation statistical tools for analyzing interdependencies across the life course. *Advances in Life Course Research*, *41*, 100249. https://doi.org/10.1016/j.alcr.2018.10.002

Nicassio, P. M., Mendlowitz, D. R., Fussell, J. J., & Petras, L. (1985). The phenomenology of the pre-sleep state: The development of the pre-sleep arousal scale. *Behaviour Research and Therapy*, *23*(3), 263–271. https://doi.org/10.1016/0005-7967(85)90004-X

Ong, J. C., Arnedt, T., Kalmbach, D. A., & Smith, M. T. (2021). Insomnia diagnosis, assessment, and evaluation. In M. Krieger, T. Roth, C. Goldstein, W. Dement (Eds.), *Principles and practice of sleep medicine* (7th ed., Vol. 1, pp. 858-866). Elsevier.

Parsons, C. E., Zachariae, R., Landberger, C., & Young, K. S. (2021). How does cognitive behavioural therapy for insomnia work? A systematic review and meta-analysis of mediators of change. *Clinical Psychology Review*, *86*, 102027. https://doi.org/10.1016/j.cpr.2021.102027

Patel, D., Steinberg, J., & Patel, P. (2018). Insomnia in the elderly: A review. *Journal of Clinical Sleep Medicine*, *14*(6), 1017–1024. https://doi.org/10.5664/jcsm.7172

Pavlova, M., Berg, O., Gleason, R., Walker, F., Roberts, S., & Regestein, Q. (2001). Self-reported hyperarousal traits among insomnia patients. *Journal of Psychosomatic Research*, *51*(2), 435-441. https://doi.org/10.1016/S0022-3999(01)00189-1

Perlis, M. L., Jungquist, C., Smith, M. T., & Posner, D. (2005). *Cognitive behavioral treatment of insomnia: A session-by-session guide* (Vol. 1). Springer Science & Business Media.

Perlis, M. L. (2017, October 21-22). *Principles and Practice of Cognitive Behavioral Therapy for Insomnia* [clinical conference]. University of Pennsylvania School of Medicine,

Philadelphia, PA, United States.

Perlis, M. L., Ellis, J. G., Spiegelhalder, K., & Riemann, D. (2021). Etiology and

pathophysiology of insomnia. In M. Krieger, T. Roth, C. Goldstein, W. Dement (Eds.),

Principles and Practice of Sleep Medicine (7th ed., Vol. 1, pp. 835-849). Elsevier.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2012). Latent Growth

Curve Modeling. *Structural Equation Modeling (2nd Ed.): Foundations and Extensions*,

155–180. https://doi.org/10.4135/9781452226576.n8

Qaseem, A., Kansagara, D., Forciea, M. A., Cooke, M., Denberg, T. D., Barry, M. J., Boyd, C.,

Chow, R. D., Fitterman, N., Harris, R. P., Humphrey, L. L., Manaker, S., McLean, R.,

Mir, T. P., Schünemann, H. J., Vijan, S., & Wilt, T. (2016). Management of chronic

insomnia disorder in adults: A clinical practice guideline from the American college of

physicians. *Annals of Internal Medicine*, *165*(2), 125–133. https://doi.org/10.7326/M15-

2175

R Core Team (2021). R: A language and environment for statistical computing. *R Foundation for*

Statistical Computing. Vienna, Austria. https://www.R-project.org/

Revelle, W. (2021) Psych: Procedures for personality and psychological research. Northwestern

University, Evanston, Illinois, USA. https://CRAN.R-project.org/package=psych Version

= 2.1.6.

Riedel, B. W., & Lichstein, K. L. (2000). Insomnia and daytime functioning. *Sleep Medicine*

Reviews, 4(3), 277-298. https://doi.org/10.1053/smrv.1999.0074

Rogers, N. L., Dorrian, J., & Dinges, D. F. (2003). Sleep, waking and neurobehavioural

performance. *Frontiers in Bioscience*, *8*(6), s1056-s1067.

Rosseel, Y., Jorgensen, T. D., Rockwood, N., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei,

V., Merkle, E., Hallquist, M., Rhemtulla, M., Katsikatsou, M., Barendse, M., Scharf, F.,

& Du, H. (2021). Latent variable analysis: Package 'lavaan'.Version 0.6-8. Available at

https://lavaan.ugent.be

Ruiter, M. E., DeCoster, J., Jacobs, L., & Lichstein, K. L. (2010). Sleep disorders in African

Americans and Caucasian Americans: A meta-analysis. *Behavioral Sleep Medicine*, *8*(4),

246–259. https://doi.org/10.1080/15402002.2010.509251

Rush, A. J., Giles, D. E., Schlesser, M. A., Fulton, C. L., Weissenburger, J., & Burns, C. (1986).

The inventory for depressive symptomatology (IDS): preliminary findings. *Psychiatry*

research, *18*(1), 65-87

Sadikaj, G., Wright, A. G. C., Dunkley, D. M., Zuroff, D. C., & Moskowitz, D. S. (2019).

Multilevel structural equation modeling for intensive longitudinal data: A practical guide

for personality researchers. *Psyarxiv*.

Sateia, M. J., Buysse, D. J., Krystal, A. D., Neubauer, D. N., & Heald, J. L. (2017). Clinical

practice guideline for the pharmacologic treatment of chronic insomnia in adults: an

American Academy of Sleep Medicine clinical practice guideline. *Journal of Clinical*

Sleep Medicine, *13*(2), 307-349. https://doi.org/10.5664/jcsm.6470

Scheidel, C. (2019). Estimating SEM using the lavaan package in R.

https://tutorials.methodsconsultants.com/posts/structural-equation-models-using-the-

lavaan-package-in-r/

Schutte-Rodin, S. L., Broch, L., Buysee, D., Dorsey, C., & Sateia, M. (2008). Clinical guideline

for the evaluation and management of chronic insomnia in adults. *Journal of Clinical*

Sleep Medicine, *4*(5), 487–504. https://doi.org/10.5664/jcsm.27286

Schwartz, D. R., & Carney, C. E. (2012). Mediators of cognitive-behavioral therapy for

insomnia: A review of randomized controlled trials and secondary analysis studies. *Clinical Psychology Review*, *32*(7), 664–675. https://doi.org/10.1016/j.cpr.2012.06.006

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Sage.

Shaw, R., McKenzie, S., Taylor, T., Olafiranye, O., Boutin-Foster, C., Ogedegbe, G., & Jean-Louis, G. (2012). Beliefs and attitudes toward obstructive sleep apnea evaluation and treatment among Blacks. *J Natl Med Assoc*, *104*(0), 510–519. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf

Shi, D., Song, H., DiStefano, C., Maydeu-Olivares, A., McDaniel, H. L., & Jiang, Z. (2019). Evaluating factorial invariance: An interval estimation approach using Bayesian structural equation modeling. *Multivariate behavioral research*, *54*(2), 224-245. https://doi.org/10.1080/00273171.2018.1514484

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, *78*(3), 298–311. https://doi.org/10.1037/a0019247

Smith, C. E., & Cribbie, R. A. (2013). Multiplicity control in structural equation modeling: Incorporating parameter dependencies. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(1), 79-85. https://doi.org/10.1080/10705511.2013.742385

Sofi, F., Cesari, F., Casini, A., Macchi, C., Abbate, R., & Gensini, G. F. (2014). Insomnia and risk of cardiovascular disease: A meta-analysis. *European Journal of Preventive Cardiology*, *21*(1), 57–64. https://doi.org/10.1177/2047487312460020

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). Manual for the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologists Press

Spielman, A., Saskin, P., & Thorpy, M. J. (1987). Treatment of chronic insomnia by restriction

    of time in bed. *Sleep*, *10*(1), 45–56. https://doi.org/10.1093/sleep/10.1.45

Tarlov, A. R., Ware, J. E., Greenfield, S., Nelson, E. C., Perrin, E., & Zubkoff, M. (1989). The

    Medical Outcomes Study: An application of methods for monitoring the results of

    medical care. *Journal of the American Medical Association*, *262*(7), 925-930.

Tomfohr-Madsen, L., Madsen, J. W., Bonneville, D., Virani, S., Plourde, V., Barlow, K. M., ...

    & Brooks, B. L. (2020). A pilot randomized controlled trial of cognitive-behavioral

    therapy for insomnia in adolescents with persistent postconcussion symptoms. *The*

    Journal of Head Trauma Rehabilitation, *35*(2), E103-E112.

    https://doi.org/*10.1097/HTR.0000000000000504*

Trauer, J. M., Qian, M. Y., Doyle, J. S., Rajaratnam, S. M., & Cunnington, D. (2015). Cognitive

    behavioral therapy for chronic insomnia: A systematic review and meta-analysis. *Annals*

    of Internal Medicine, 163(3), 191-204. https://doi.org/10.7326/M14-2841

Van Houdenhove, L., Buyse, B., Gabriëls, L., & Van Den Bergh, O. (2011). Treating primary

    insomnia: Clinical effectiveness and predictors of outcomes on sleep, daytime function and

    health-related quality of life. *Journal of Clinical Psychology in Medical Settings*, *18*(3),

    312–321. https://doi.org/10.1007/s10880-011-9250-7

van Straten, A., van der Zweerde, T., Kleiboer, A., Cuijpers, P., Morin, C. M., & Lancee, J.

    (2018). Cognitive and behavioral therapies in the treatment of insomnia: A meta-analysis.

    Sleep Medicine Reviews, *38*, 3–16. https://doi.org/10.1016/j.smrv.2017.02.001

Vincent, N., & Lewycky, S. (2009). Logging on for better sleep: RCT of the effectiveness of

    online treatment for insomnia. *Sleep*, *32*(6), 807–815.

    https://doi.org/10.1093/sleep/32.6.807

Wickham, H. *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York, 2016.

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software, 40*(1), 1-29. http://www.jstatsoft.org/v40/i01

Wickham, H. (2021). tidyr: Tidy messy data. R package version 1.1.3. https://CRAN.R-project.org/package=tidyr

Wickham, H., François, R., Henry, L., Müller, K. (2021). dplyr: A grammar of data manipulation. R package version 1.0.7. https://CRAN.R-project.org/package=dplyr

Wickwire, E. M., Tom, S. E., Scharf, S. M., Vadlamani, A., Bulatao, I. G., & Albrecht, J. S. (2019). Untreated insomnia increases all-cause health care utilization and costs among Medicare beneficiaries. *Sleep*, *42*(4), 1–8. https://doi.org/10.1093/sleep/zsz007

Winzeler, K., Voellmin, A., Schäfer, V., Meyer, A. H., Cajochen, C., Wilhelm, F. H., & Bader, K. (2014). Daily stress, presleep arousal, and sleep in healthy young women: A daily life computerized sleep diary and actigraphy study. *Sleep Medicine*, *15*(3), 359–366. https://doi.org/10.1016/j.sleep.2013.09.027

Wohlgemuth, W. K., Edinger, J. D., Fins, A. I., & Sullivan, R. J. (1999). How many nights are enough? The short-term stability of sleep parameters in elderly insomniacs and normal sleepers. *Psychophysiology*, *36*(2), 233–244. https://doi.org/10.1017/S0048577299971214

Wu, R., Bao, J., Zhang, C., Deng, J., & Long, C. (2006). Comparison of sleep condition and sleep-related psychological activity after cognitive-behavior and pharmacological therapy for chronic insomnia. *Psychotherapy and Psychosomatics*, *75*(4), 220–228. https://doi.org/10.1159/000092892

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory

factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2), 79-94.

    https://doi.org/10.20982/tqmp.09.2.p079

Zhu, M., Hong, R. H., Yang, T., Yang, X., Wang, X., Liu, J., Murphy, J. K., Michalak, E. E.,

    Wang, Z., Yatham, L. N., Chen, J., & Lam, R. W. (2021). The efficacy of measurement-

    based care for dperessive disorders: Systematic review and meta-analysis of randomized

    controlled trials. *Journal of Clinical Psychiatry*, *82*(5), e1–e11.

    https://doi.org/10.1017/S1121189X00004875