

Assessment methodologies in sleep medicine clinical trials

Clin. Invest. (2013) 3(8), 791–800

Clinical trials in sleep medicine cover a wide range of sleep–wake problems, and accordingly the selection of outcome measures in sleep medicine clinical trials needs to be tailored to the specific disorder under examination. This review describes the measures most commonly used in sleep medicine clinical trials, weighing the relative merits of a self-report questionnaire versus a physiologic test as the *a priori* primary outcome measures.

Keywords: actigraphy • clinical trial • outcome measures • polysomnogram • psychometrics • sleep diary • sleep medicine

This paper is intended for an audience that is interested in the design of sleep medicine clinical trials, yet uninitiated to the topic. Herein, we describe some of the most commonly used measurement instruments in sleep medicine clinical trials, and provide guidance on how to select measurement instruments to meet the investigator's goals in the design of a sleep medicine clinical trial.

The design of clinical trials in sleep medicine, like all clinical trials, begins with defining the diagnostic problem in which an intervention will be tested. In the arena of sleep medicine, defining the diagnosis of interest is complicated by the existence of three nosologic systems, including the second edition of the International Classification of Sleep Disorders (ICSD-2) of The American Academy of Sleep Medicine (AASM), the Diagnostic and Statistical Manual of Mental Disorders (DSM) of the American Psychiatric Association and the International Classification of Diseases (ICD-9).

The ICSD-2 tends to split (as opposed to group) diagnostic categories, resulting in eight major categories: insomnia, sleep-related breathing disorders, hypersomnias of central origin, circadian rhythms disorders, parasomnias, sleep-related movement disorders, isolated symptoms and normal variants as well as others [1]. These major categories are then split into over 70 specific diagnoses.

The DSM is the guide for psychiatric diagnoses, including sleep disorders, as used in clinical and research settings [2]. The fourth edition of the DSM includes a much simpler sleep nosology, as compared with the ICSD-2, and the DSM has been published as a revised edition titled the DSM-V in May 2013. The DSM-V contains revisions in its own sleep nosology [3].

Finally, the ICD-9 names a smaller number of sleep diagnoses than the ICSD-2, but the ICD-9 sleep diagnoses are spread out over the pulmonary, neurological, psychiatric and symptomatic portions of the text.

This paper will not consider the differences between the ICSD-2, DSM and ICD-9 nosologies, and instead organize sleep disorders by four major symptomatic presentations: the disorders of excessive daytime sleepiness (EDS), the insomnias, the circadian rhythms sleep disorders (CRSD) and the parasomnias. The measurement instruments that make diagnoses and establish inclusion and exclusion criteria for clinical trials may also be used to measure change in response to treatment. However, in this paper we will describe commonly used measurement instruments

Elena A Wood^{*1} & W Vaughn McCall²

¹Center for Telehealth, Medical College of Georgia, Georgia Regents University, 1120 15th Street, Building EA, Augusta, GA 30912 1655, USA

²Department of Psychiatry & Health Behavior, Medical College of Georgia, Georgia Regents University, 997 St Sebastian Way, Augusta, GA 30912, USA

*Author for correspondence:

Tel.: +1 706 721 8849

E-mail: eawood@gru.edu

only in regard to their application to detect treatment effects, not in making diagnoses. Furthermore, we limit our scope to adults.

This paper is organized as follows:

- A brief description of each of the four major categories of sleep disorders;
- A description of commonly used measurement instruments in that category;
- An illustrative example of a clinical trial that successfully used that instrument to differentiate among treatment effects.

When possible, the authors picked clinical trials that had two arms, but when that was not possible, we picked single-armed trials that show pre–post intervention differences.

Daytime sleepiness

Symptoms of EDS may be endorsed by nearly half of the adult population, but the percentage of the adult population that meets the definition of a hypersomnia disorder is between 0.5 and 1% [4]. EDS disorders describe individuals with a degree of sleepiness during the patient's waking hours that is of sufficient intensity that it interferes with normal functioning and causes distress and concern. The most common causes of EDS are sleep deprivation, obstructive sleep apnea (OSA), narcolepsy, side effects of medications, general medical conditions and idiopathic hypersomnias.

OSA is a condition of repeated upper airway complete or partial obstructions, producing oxygen desaturation and arousals from sleep. OSA frequently presents with EDS. A common treatment of OSA is continuous positive airway pressure (CPAP) via a nasal mask during sleep to prevent pharyngeal airway collapse [5]. The severity of OSA is commonly tracked by the apnea-hypopneas index (AHI), which is defined as the number of complete breathing pauses (apneas) and partial breathing pauses (hypopneas) across the night, divided by the number of hours of sleep.

Narcolepsy is a neurological disorder characterized by EDS with cataplexy, hypnagogic hallucinations and sleep paralysis. The most common symptom of narcolepsy is EDS, preceding other symptoms by years. Commonly, symptoms first present in young adults, and narcolepsy has a chronic course that can be treated but not cured. Many cases of narcolepsy are caused by orexin/hypocretin deficiency [6].

EDS can be assessed by pathophysiologic finding in a laboratory, by patient self-report or observation by others. There are objective and subjective validated self-reported methods to evaluate the degree of sleepiness and alertness.

■ Physiologic measurement instruments commonly used in clinical trials of EDS

Polysomnography (PSG) is widely considered the 'gold standard' test for objective measurement of sleep and wake. PSG includes assessments of time spent awake, time spent asleep and quantification of sleep stages, with continuous recordings of the electroencephalogram (EEG), eye movements (EOG) and chin electromyography (EMG). Clinical trials of OSA further expand the PSG array to include ECG, measures of respiratory airflow, respiratory effort, arterial oxygen saturation and leg EMG. PSG testing has historically been conducted in a dedicated sleep laboratory, although recent developments include a shift to portable testing in the patient's home. The equipment costs for PSG are substantial, and there are also high labor costs in the scoring and collation of the resulting data. These costs are pushed higher if the study is conducted in a sleep laboratory while attended by professional staff. Improvement in OSA is reflected in reductions in the AHI, as well as improvement in the degree of disruption of sleep architecture, and the number and depth of arterial desaturations.

The following clinical trial illustrates the appropriate utilization of PSG in an intervention that intended to show the efficacy of CPAP for OSA. The purpose of this clinical trial was to compare nasal CPAP, oxygen delivered by nasal cannula and placebo (air delivered by nasal cannula) in adult patients with mild OSA (AHI >5 and <15) and EDS [5,7]. The baseline evaluation was followed by three treatment evaluation points, separated by approximately 1 month. Assessments included nocturnal PSG and a battery of daytime measurements. Eight subjects were randomized to receive nasal air (placebo) versus nasal oxygen for a month. In the following month subjects crossed over and received nocturnal treatment with the other gas. During the last month, all subjects received nasal CPAP treatment. In the analysis, all subjects served as their own control (baseline). Outcome measures included the AHI, the number of arterial oxygen desaturation events that represented a >4% decline in SaO_2 , and the degree of disruption of EEG sleep architecture (as reflected in increases in lighter sleep stages, reduced sleep efficiency [SE] and increased number of arousals). Mean AHI was significantly lower after CPAP treatment (3.0 ± 0.9) compared with baseline (20.5 ± 4.8 ; $p < 0.05$), to placebo (22.1 ± 5.7 ; $p < 0.05$) and to oxygen (16.8 ± 3.2 ; $p = 0.09$). The number of arterial desaturation events were fewer after CPAP treatment (32.6 ± 11.1) as compared with air (208.1 ± 51.7 ; $p < 0.01$) and baseline (168.9 ± 39.2 ; $p < 0.05$), but not different when compared with oxygen (29.4 ± 8.2). Analysis of EEG sleep architecture showed fewer EEG arousals after CPAP (70.8 ± 16.6) as compared with

baseline (202.6 ± 66.8 ; $p = 0.08$). Thus PSG succeeded in discerning differences in treatment efficacy for OSA.

While PSG may be used to show improvement in OSA, it does not directly provide objective measurement of EDS itself. Instead, the Multiple Sleep Latency Test (MSLT) and Maintenance of Wakefulness Test (MWT) are conducted during the patient's usual wake period to directly assess physiologic measures of sleepiness.

The MSLT is used to assess a subject's tendency to fall asleep after a normal night's sleep, during four or five 20 min naps throughout the day, spaced at 2 h intervals [8]. The MSLT is preceded by 2 weeks of normal nocturnal sleep. The MSLT usually follows a night-time PSG study. EEG, EOG and chin EMG are recorded during each nap. Unlike the PSG, which can conceivably be conducted in the patient's home, the MSLT is conducted in a sleep laboratory while the patient is attended to by a technician. During each nap, the patient lies in bed while wearing casual clothing and the room is dark and quiet. The latency in minutes from lights out to the onset of EEG-defined sleep is calculated for each nap. If the subject does not fall asleep during the 20 min, then the session is concluded, with a sleep latency score of 20 min for that nap. A short mean sleep latency across all naps (e.g., <7 min) is not specific for any particular sleep disorder, but does indicate excessive sleepiness. The presence of two or more sleep-onset rapid-eye-movement periods is strongly supportive of narcolepsy [9]. The MSLT is also recommended by the AASM to be used for treatment monitoring for patients with OSA that requires CPAP [10]. The MSLT has a significant limitation: it is only valid after a minimum of 6 h sleep, and the MSLT sensitivity is decreased in the elderly population [11]. The MSLT is expensive, time-consuming, based on 1 day only and is conducted within a sleep laboratory. The MSLT is a well-validated and extensively published objective measure and has been indicated for use in 143 publications reported on clinical trials; 22 of them involved patients with narcolepsy.

In one narcolepsy study, the stimulant efficacy of modafinil was investigated in a double-blind, placebo-controlled, randomized trial in narcolepsy patients with cataplexy, naive to EDS drug therapy [12]. Modafinil 100 mg ($n = 24$) or placebo ($n = 20$) was administered for 4 weeks in age-matching pairs. The pair-wise analysis showed significant increase of 6.6 point score of the MSLT in the experimental group ($p < 0.001$). There were no significant changes in the placebo group. Thus the MSLT is suitable for detecting physiologic differences among the alerting effects of stimulant medications.

The MWT is a validated test to measure a subject's ability to remain awake during the day [13]. It is a modification of the MSLT using the same method, except that at each of the scheduled naps subjects are told

to stay awake (not 'fall asleep' as in the MSLT), and each session consists of four 40 min tests spread 2 h apart. The subject will be placed in a private, darkened quiet room with temperature adjusted to personal comfort level and asked to stay awake. As is the case with MSLT, the requirement of 2 weeks normal night-time sleep is required prior to the test. The primary measure for the MWT is sleep latency, and if the patient fails to fall asleep on any given nap attempt, then that nap is scored as 40 min. A mean latency for four sessions that is <8 min is considered abnormal. Staying awake during all four sessions for 40 min is considered normal, whereas a mean of between 8 and 40 is uncertain [10]. Doghramji *et al.* presented normative data for the MWT and guidelines for optimal recording and scoring of the MWT [14]. The AASM recommends using the MWT to assess response-to-treatment in subjects with narcolepsy or idiopathic hypersomnia [10]. Like the MSLT, the MWT test is expensive and requires a laboratory. The MWT has been used since 1996 in 58 clinical trials, 20 of which are narcolepsy trials.

In one example, sodium oxybate was examined in a double-blind, placebo-controlled clinical trial in patients with narcolepsy [15]. All 278 patients were previously taking modafinil 200–600 mg daily for EDS treatment. At baseline, patients had PSG and MWT studies, followed by randomization into four groups: placebo, sodium oxybate, modafinil or sodium oxybate plus modafinil. Sodium oxybate 6 mg was administered nightly for 4 weeks and 9 mg for a further 4 weeks. The MWT was the primary outcome measure, repeated after 4 and 8 weeks from the baseline [16]. In the placebo group the mean average daytime sleep latency on the MWT decreased from 9.7 min at baseline to 6.9 min at the 8 week data point ($p < 0.001$). The sodium oxybate plus modafinil group had an increase in daytime sleep latency from 10.4 to 13.5 min ($p < 0.001$). There were no significant differences in the two other groups. Like the MSLT, the MWT is suitable for detecting physiologic differences among the alerting effects of stimulant medications.

■ Self-report measures commonly used in clinical trials of EDS

The Epworth Sleepiness Scale (ESS) measures EDS on the basis of the patient's self-perception on eight different daytime situations in which they could doze off. Since 1996, PubMed notes 372 publications reporting results of clinical trials that involved the ESS, 147 of them are on OSA and 39 publications mentioned use of the ESS in narcolepsy clinical trials. The ESS scores 0–3 for each of the eight situations in which the patient might fall asleep (0: 'would never doze' and 3: 'high chance of dozing'), and the scores for the eight situations are totalled (full

range 0–24), with a score of 10 or more indicating abnormal daytime sleepiness. The ESS is usually scored on the basis of the patient's recollection of sleepiness in the last week, and hence is not suitable for 'spot' measurements of subjective sleepiness. The ESS differentiates between normal alertness versus various sleep disorders (OSA, narcolepsy and idiopathic hypersomnia) [17]. In patients with OSA, the ESS score strongly correlates with the apnea–hypopnea index and the minimum SaO_2 . The ESS has high levels of internal consistency (Cronbach's alpha is 0.88) and a number of studies supported high validity and reliability [18]. The ESS had high sensitivity for OSA and has been recommended as a screening tool for these conditions [19]. The ESS has been widely used to evaluate treatment response in OSA.

For example, a novel nasal-valve expiratory positive airway pressure (EPAP) device was studied with the ESS as an outcome measure [20], in a prospective, multicenter, sham-controlled, parallel-group, randomized, double-blind clinical trial [21,22]. The study enrolled 250 newly diagnosed adult patients with OSA and randomized them into the experimental nasal-valve EPAP-device group or a similar appearing sham device. The ESS score and PSG were recorded at baseline and again at 3 months. Data were available for 229 subjects (119 EPAP and 110 sham), analyzed using intention-to-treat. AHI was reduced to $<10/\text{h}$ in 62.0% of the patients in the EPAP arm compared with 27.2% in the sham group ($p < 0.001$). The ESS score decreased more in the EPAP arm (from 9.9 to 7.2), than in the sham arm (from 9.6 to 8.3) based on intention-to-treat analysis results ($p = 0.04$). Thus, the study concluded that the nasal EPAP effectively reduced the AHI with a corresponding improvement in self-reported sleepiness compared with the sham device group, as measured by the ESS.

The Stanford Sleepiness Scale (SSS) is a quick and simple self-administered instrument to evaluate how alert or sleepy a person is at any given moment in time. Subjects choose one out of seven proposed answers best describing 'how do you feel right now?', thus allowing for the observation of changes in alertness across the day. The items on the SSS are scored 1–7, with a score of 1 indicating that the patient is feeling active, alert and wide awake, while 7 indicates the person is almost in reverie, sleep-onset soon and has lost the struggle to remain awake [23]. The limitations of the test include a lack of specificity for any particular sleep disorders, but in turn this permits a more generic, broader applicability in a variety of clinical trial designs [24]. The SSS has been used 59 times since 1996 in clinical trials to evaluate the general alertness/sleepiness in different settings (sleep deprivation, intensive care unit, alcohol intoxication, and so forth). Often ESS and SSS are used together to assess outcomes.

This is the case in the following study, which aimed to evaluate the efficacy of mirtazapine on EDS symptoms in depressed adult patients [25]. In the study, 42 patients were matched with 32 healthy controls at the baseline (the SSS mean score was 4.1 ± 0.4 in the patient group and 2.0 ± 0.3 in the control group; $p < 0.01$). The 16 patients received a 58-day therapy of 30-mg mirtazapine daily, 30 min before bedtime. The treatment effect of mirtazapine on the ESS and SSS scores was similar. The SSS mean score on day 58 in the mirtazapine group was 3.5 ± 1.3 ; that is, 28.6% lower than at the baseline ($p = 0.001$).

Apart from the subjective report of symptoms of EDS, some investigators are interested in the impact of EDS on quality of life and daily function. The Functional Outcomes of Sleep Questionnaire (FOSQ) was developed to measure the impact of disorders of EDS on daily life activities (activity, vigilance, sex, productivity and social life) [26], and the original version has 30 items. The internal consistency was reported to be excellent (Cronbach's alpha is 0.90). In 2009, a short version of the FOSQ was presented and psychometric properties were evaluated [27]. The FOSQ-10 internal consistency was Cronbach's alpha = 0.87 compared with 0.95 for the long version. The versions of the FOSQ were strongly correlated ($r = 0.96$; $p < 0.0001$). The FOSQ-10 was demonstrated to distinguish between OSA (12.5 ± 3.2) and normal (17.8 ± 3.1) groups ($p < 0.0001$). There are 44 publications to date on clinical trials using the FOSQ, and 33 of them are related to OSA.

As one example, Weaver *et al.* conducted a multi-site, double-blind, randomized, placebo-controlled, parallel-group clinical trial to evaluate functional status of patients with mild and moderate OSA (AHI 5–30 per hour) in CPAP versus sham treatment [28]. Patients naive to CPAP, with an ESS score of greater than 10, were randomized into two arms. Subjects completed the FOSQ weekly during 8 weeks of active or sham CPAP treatment. Objectively, the AHI changes were greater ($p = 0.0001$) in the intervention group (-11.9) compared to the sham (-2.4) group. In the modified intention-to-treat analysis of 223 patients (113 active CPAP, 110 sham CPAP), the adjusted mean of the total FOSQ score increased by 0.89 in the active treatment group, and in the sham device group decreased by 0.06 ($p = 0.006$). Thus the FOSQ is a useful instrument in detecting changes in quality of life in sleep apnea patients.

Insomnia

Insomnia occurs in both acute, transient forms as well as chronic forms. In either instance, insomnia is characterized by the patient's dissatisfaction with the time required to fall asleep, and/or excessive wake time after initially falling asleep, and/or complaints of poor sleep

quality, all associated with complaints of daytime irritability, fatigue, concentration problems, and so forth. Chronic insomnia is more common than chronic EDS, with as many as 10% of Americans reporting insomnia [29]. Psychiatric illness is associated with about half of the cases of chronic insomnia, with the remainder of cases associated with medical illness, medication side effects, restless leg syndrome or idiopathic insomnia. In routine clinical practice, assessment and treatment of insomnia does not require PSG testing, but in insomnia clinical trials PSG is often used to clarify that the sample is free of other primary sleep disorders such as OSA. Research diagnostic criteria have been proposed for insomnia [30]. The clinical guideline for the evaluation and management of chronic insomnia in the adult population states that, regardless of the type of therapy, primary outcome measures should be sleep quality and quantity, and insomnia-related daytime impairment [31].

■ **Physiologic measurement instruments commonly used in clinical trials of insomnia:**
PSG & actigraphy

Full PSG, including measurement of breathing and leg movements, is commonly used during initial characterization of an insomnia sample prior to inclusion in a clinical trial. After randomization, either PSG is not used to detect the effect of the intervention, or repeated PSG measurements are made, but only using EEG, EOG and chin EMG.

For example, a multicenter, randomized, placebo-controlled trial used PSG as an objective measure of the efficacy of eszopiclone as compared with placebo and zolpidem in the treatment of primary insomnia [32]. A group of 65 adult patients received placebo for 2 nights, followed by randomization to placebo, zolpidem 10 mg or eszopiclone 1, 2, 2.5 or 3 mg. PSGs were conducted during screening up to 3 nights in the sleep laboratory, and on 2 consecutive nights during treatment. The primary outcome was 'latency to persistent sleep' (LPS, defined as the time from onset of the PSG recording to the start of 10 continuous minutes of any stage of sleep). LPS was significantly reduced compared with placebo for all active treatments. The LPS median was 29.0 min in placebo, 16.8 min in eszopiclone 1 mg, 16.5 min for 2 mg, 13.8 min for 2.5 mg and 13.1 min for 3 mg and zolpidem 10 mg ($p < 0.05$). Thus, in this study the PSG was sensitive in detecting the effect of active drug versus placebo in patients with chronic insomnia.

Actigraphy (ACT) has been used to study sleep for the last four decades. The earliest reports on ACT included comparisons of the ACT's performance versus PSG [33]. ACT is the continuous measurement of body movement via the electronic recording of changes

in velocity. It is most often recorded with a small sensor that has the appearance of a wrist watch, and the device is worn on the nondominant wrist. It measures changes in velocity in any direction, and records at a rate of up to 32 times per second. The number of body movements is stored in epoch lengths of 15, 30, 60 or 120 s as designated by the operator. Battery life will allow data collection to continue for weeks at a time before the research participant returns to the laboratory to download data from the ACT watch. Data are displayed and analyzed according to the time-epoch length that was previously set. The data can be analyzed by different thresholds of sensitivity. The advantages of the ACT include long-term monitoring of the patient's rest-activity cycle in their natural environment. Disadvantages of ACT include the inability to identify sleep stages. The accuracy of ACT in identifying sleep versus wakefulness has been discussed in the last decade [34–39]. ACT often underestimates sleep onset latency (SOL) and overestimates total sleep time (TST) and SE as compared with PSG [34]. In addition, ACT may not be accurate for sleep evaluation in patients with movement disorders due to underestimation of TST. Initial startup costs for hardware and software can be substantial, but after the initial cost outlay, the remaining costs are the labor involved in downloads and data analysis. In 2007, the AASM published clinical recommendations on ACT use [36]. ACT was recommended to be used as a diagnostic tool on a healthy adult population and in the initial evaluation of patients with sleep disorders, and as an outcome measure of response-to-treatment. When PSG is not available, ACT can be used to estimate TST in patients with OSA. ACT is useful in monitoring night-to-night variability [35].

The utility of ACT in insomnia is illustrated in a study evaluating mindfulness-based stress reduction (MBSR) therapy for chronic primary insomnia patients. In this study, ACT found a decrease in SOL of 8.9 min between baseline and the 8-week measurement point [40]. In total, 30 adult patients were randomized into the MBSR group ($n = 20$) and the pharmacotherapy control group ($n = 10$). MBSR intervention included classes, home practice and a daylong retreat on different meditation techniques. Participants wore ACT on the nondominant wrist for 14 days prior to the interventions and during the final 2 weeks of the interventions. The following parameters measured by ACT were used in the analysis in tandem with sleep diary data: TST, SOL, wake after sleep onset (i.e., SOL) and SE. The ACT SOL decreased significantly from 34.2 ± 28.3 min at the baseline to 25.6 ± 20.00 min at the end of MBSR treatment ($p = 0.04$). ACT measurements showed a significant improvement in TST (6.40 ± 0.60 h at baseline vs 6.9 ± 0.6 at end-of-treatment; $p < 0.05$) and

SE ($75.5 \pm 12.5\%$ vs $83.5 \pm 6.4\%$; $p < 0.01$) in the pharmacotherapy arm.

■ Self-report measures commonly used in clinical trials of insomnia

A sleep diary (sleep log) is a widely used, simple and inexpensive method to collect data about a subject's sleep pattern over a period of time (usually 2 weeks). PubMed reports the use of sleep diaries in 108 clinical trials since 1996, 54 of which related to insomnia. Usually, the individual is asked to collect information about bed time, sleep latency, number of waking episodes, awakening time and mood, taking alcohol, caffeine and naps. A standardized sleep diary, termed the Consensus Sleep Diary (CSD), is under development based on 25 insomnia experts' opinions and testing in patient focus groups [41]. The CSD includes 9 questions: the time getting into bed, the time when and for how long the subject tried to fall asleep, number and duration of awakening, the time of final awakening and getting out of bed, perception of sleep quality and comments. The CSD is formatted to one page and includes 1 week of data. Sleep diaries are helpful with diagnosing and assessing treatment in patients with insomnia and circadian sleep disorders, and are often used in conjunction with ACT in insomnia research. Sleep diaries have been compared with PSG and ACT in depressed insomniacs [42]. While PSG and ACT sleep parameters had positive correlations, significant differences were observed between sleep diaries and PSG. This study suggests that ACT is a better reflection of PSG sleep than sleep diaries. In-laboratory versus at-home comparisons between sleep diaries and ACT parameters found significant differences in actigraphic data that were not reflected in sleep diaries self-evaluations [43]. ACT measurements found significantly increased sleep time and decreased wake time in sleep laboratory settings while sleep diaries demonstrated a significantly increased number of awakenings in the same settings. One limitation of a sleep diary is noncompliance with daily recording due to patient forgetfulness. This limitation can be largely solved by moving to technology-based formats, as was the case in the following example.

A randomized, multisite, double-blind, placebo-controlled trial of eszopiclone in elderly insomniacs used electronic hand-held sleep-wake diaries to record self-reported sleep-wake activity twice a day (6–10 am and 20:00–23:45 pm) [44]. The primary outcome measure was subject-reported TST (sTST), and secondary efficacy measures were subject-reported sleep latency (sSL) and subject-reported SOL. Morning parameters in electronic sleep diaries included: sSL, sTST, subject-reported SOL, number of awakenings, quality of sleep

and depth of sleep. The evening diary assessed the following: number and length of naps, daytime alertness, ability to function and concentrate and sense of physical well-being. The sTST improved over 12 weeks in the eszopiclone group (baseline: 360.0 min) compared to the placebo group (baseline: 297.9 min) by 63.2 min ($p \leq 0.001$). The sSL decreased ($p = 0.0014$) in the experimental group (mean decrease: 24.6 min) compared with the placebo group (19.9 min).

The Insomnia Severity Index (ISI) was designed to screen, evaluate severity and monitor treatment outcomes in patients with insomnia [45,46]. It is a brief (seven questions on a 0–4 scale), reliable and validated self-reported questionnaire addressing severity of sleep onset, waking problems, sleep maintenance and dissatisfaction, daytime functioning, how noticeable the sleep problem is to others and the distress due to sleep problems. The ISI has excellent internal consistency (Cronbach's alpha of 0.90), and is well-correlated with sleep diaries, PSG and interviews [46]. A six-point reduction in the ISI score represents a clinically meaningful improvement. Limitations of the ISI include the inability to differentiate primary insomnia and other psychiatric or medical diseases-caused insomnia. This instrument has been described in 44 clinical trials in PubMed since 1996.

As one example, a randomized controlled trial was conducted to evaluate effectiveness of cognitive behavioral therapy (CBT) for insomnia (CBT-I) [47], with 151 adults randomized into either CBT-I versus control (wait-list) groups. ISI was the primary outcome measure completed at baseline and follow up. In the experimental group, the ISI decreased from 16.4 ± 4.6 at baseline to 10.8 ± 5.9 at follow up ($p = 0.000$) compared with control from 17.1 ± 5.4 at baseline to 16.2 ± 5.0 at follow up ($p = 0.077$). Thus, ISI indicated significant differences between time (baseline vs follow up) and group (experimental vs control) in this study.

The Pittsburgh Sleep Quality Index (PSQI) was developed to measure sleep quality. It is self-administered with 17 items in seven clinical domains related to sleep difficulties such as sleep quality, latency, duration, habitual SE, disturbances, sleep medications use and daytime functioning. The self-rating answers score on 0–3 scale (0 positive and 3 negative extremes on the Likert Scale). These seven scores sum into a global PSQI score, with a score of 5 or more indicating poor sleepers with 89.6% sensitivity and 86.5% specificity [48]. In patients with primary insomnia this score resulted in sensitivity of 98.7% and specificity of 84.4% in separating insomnia patients and good sleepers [49]. The PSQI has internal consistency of Cronbach's alpha = 0.83. The PSQI has been reported in 267 publications since 1996 and has been used

in clinical trials, 70 of which are on patients with insomnia.

PSQI was the primary outcome measure in a randomized controlled trial comparing supported self-help with standard of care [50]. The participants were 193 individuals 55–87 years old with chronic insomnia symptoms according to the fourth addition of the DSM PSQI scores >5. The intervention included six self-help booklets, designed upon psychoeducation and health education and a telephone helpline to address any concerns about the material covered in the booklets. Measurement points were at baseline, after treatment and at 3 and 6 months after treatment. The control group also received a summary sheet about sleep hygiene measures after baseline assessment by mail. At the post-treatment assessment, the intervention group reported significantly higher PSQI (adjusted mean difference 2.02; $p < 0.001$). Thus, PSQI was sensitive in detecting the differences in perceived sleep quality among patients receiving self-help sleep management, based on CBT, versus usual care.

The Dysfunctional Beliefs and Attitudes about Sleep (DBAS) scale is a reliable and valid self-reported measure of the most common unhelpful beliefs about sleep that are reported by patients with insomnia. The initial 30-items scale was modified into a 16-item scale (DBAS-16) [51]. The DBAS takes into account four insomnia-related factors: perceived consequences, worry/helplessness, sleep expectations and medication. The DBAS-16 has adequate internal consistency (Cronbach's $\alpha = 0.77$ – 0.79) and correlates with ISI, Beck Depression and Anxiety Inventories. A validity study has been conducted to evaluate the DBAS score between different insomnia subgroups [52]. The DBAS-16 is recommended to evaluate the potential impact of sleep-related beliefs and attitudes and to monitor therapeutic effect in insomnia. It has been studied as a moderating factor in clinical research including CBT in treatment of insomnia [53–58].

CRSD

The major characteristic of CRSDs is chronic or recurrent sleep disturbance due to misalignment between an individual's sleep pattern and the patient's desired sleep–wake schedule or societal expectations. There are six types within this category in ICSD-2: delayed sleep phase disorder (DSPD), advanced sleep phase type, irregular sleep–wake phase type, free-running disorder, jet lag and shift work disorder [1]. It is important to exclude other sleep disorders and substance abuse prior to diagnosing CRSD.

According to ICSD-2 criteria, a week of sleep monitoring with a sleep diary or ACT with a sleep diary is required for diagnosing all types of CRSD except

jet lag [1]. However, AASM Practice Parameters for clinical evaluation and treatment of CRSD recommend ACT as an evaluation tool for all CRSDs [59].

As an example, a randomized-controlled trial was conducted to compare CBT plus bright light therapy (BLT) with control wait-list in adolescents diagnosed with DSPD [60]. The experimental intervention consisted of six 45–60 min individual sessions, using morning BLT to advance participants' circadian rhythms. Measurement points were pre- and post-treatment, with 7-day sleep diaries completed online. The primary outcome measures in the study derived from the sleep diary were SOL, rise time and TST. A comparison of the control group ($n = 23$) to the experimental group ($n = 17$) found significant improvement in SOL (from 78.1 to 22.2 min in CBT plus BLT vs 78.8 to 65.3 min in the controls; $p = 0.003$). Thus, CBT plus BLT were found to be effective treatment in adolescents with DSPD as measured by sleep diaries.

ACT is used in circadian rhythm research and clinical sleep disorders centers as an objective measure in CRSD. PSG is not routinely indicated for use in diagnosing CRSD or tracking progress in CRSD clinical trials, although PSG might be used at baseline to rule out sleep disorders other than CRSD. The AASM clinical recommendations on ACT use in circadian rhythm abnormalities indicate ACT is a suitable evaluation tool in patients with advanced sleep phase syndrome, delayed sleep phase syndrome and shift work disorder [36,59], with a recommended recording duration of 1–3 weeks to characterize CRSD. ACT software allows a comparison of work day versus weekend sleep timing. ACT can be used to evaluate response to therapy for all types of CRSD.

For example, in a placebo-controlled, counter-balanced study, ACT was used to compare the addition of melatonin versus placebo to behavioral interventions in advancing circadian rhythms [61]. In total, 12 adults completed the 5-week intervention study. All participants wore ACT on the dominant wrist, recording 30 epochs at medium sensitivity. The sleep episodes were compared with sleep diary records. The sleep-onset time, wake time, and TST from the ACT database were used for analysis. Melatonin in the afternoon in addition to the gradually advancing sleep schedule produced significantly larger circadian phase advance (1.3 ± 0.7 h) compared with placebo (0.7 ± 0.7 h).

Parasomnias

Parasomnias are disruptive sleep-related disorders characterized by undesirable physical or verbal activities, emotions and dreaming, related to sleep and wake-to-sleep transition. Types of parasomnias

include nightmares, sleep terrors, sleepwalking, sleep paralysis and rapid-eye-movement behavior disorder. PSG with video recording in the sleep laboratory is considered the gold standard for evaluating patients with motor or behavioral activities during sleep, especially if the behavior occurs almost every night. Video-PSG has high specificity and variable sensitivity, but it is high cost and a time-consuming study. Other objective methods such as ACT are not specific enough for parasomnias diagnoses. Video-EEG is also important for differential diagnosis of epileptic and nonepileptic seizures [62].

Nightmares may not have any PSG correlate, but progress in the treatment of nightmares can be monitored with the Disturbing Dreams and Nightmares Severity Index (DDNSI). The DDNSI is a revised version of the Nightmare Frequency Questionnaire. The DDNSI measures the frequency and intensity of disturbing dreams and nightmares. It has good internal consistency (Cronbach's $\alpha = 0.83$). The DDNSI has five questions, answers to which can score in range 0–37, with higher scores indicating a more severe case of nightmares.

For example, in one study, the DDNSI was used to evaluate the effect of nightmares as a mediating variable for the relationship between insomnia and suicidal ideation [63]. In total, 50 individuals with insomnia and depression were assessed in a cross-sectional design with measurement of suicidal ideation and the DDNSI, finding that in a mediation analysis nightmares indeed mediate the association between insomnia symptoms and suicidal ideation.

Conclusion

Clinical trials in sleep medicine cover a wide range of sleep–wake problems, and accordingly the selection of outcome measures in sleep medicine clinical trials needs to be tailored to the specific disorder under examination. This review has not been exhaustive, but intended to describe the measures most commonly used in sleep medicine clinical trials. The investigator needs to consider the relative merits choosing a self-report questionnaire versus a physiologic test as the *a priori* primary outcome measures. Surprisingly, the self-report measures are often more sensitive to

treatment effects as compared with more expensive physiologic tests.

Still, there are a few circumstances that require a physiologic test to meaningfully show a clinically relevant outcome. For example, clinical trials in OSA require some measurement of the AHI. Even so, in this case the investigator needs to further choose between repeated in-laboratory monitoring versus home monitoring of OSA, and choose between full PSG versus a test that measures only respiratory variables but not EEG.

We recommend that full PSG, including measurements of respiration and leg movements be a part of the baseline measurement of every sleep medicine clinical trial, if suitable resources are available. A PSG baseline assessment could be omitted for an insomnia clinical trial if costs are prohibitive. Likewise, some self-reported measure of symptoms should be incorporated into every sleep medicine clinical trial. After randomization in a clinical trial, respiratory monitoring is not a necessary part of follow-up PSG, unless the condition under study is OSA or another sleep-related breathing disorder. ACT is a technology that is still in development, despite decades of use. It potentially has a role as a primary outcome measure in CRSD clinical trials, but outside of CRSD it should be relegated to the role of a secondary outcome measure.

Future perspective

The trend in sleep medicine research measurements has been moving from expensive, laboratory-based methods to unobtrusive, less expensive and more portable assessment tools allowing for the monitoring of sleep for long periods of time. While the field of sleep medicine is moving in that direction, PSG will continue to be the gold standard for validation of evolving assessment techniques.

Financial & competing interests disclosure

WV McCall is sponsored by a NIH grant (R01 MH095776-01A1). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

- A large number of psychometric and physiologic tools are available for assessing outcomes in sleep medicine clinical trials. However, very few of these measuring instruments are appropriate for evaluation of most sleep disorder interventions. Investigators should be selective in picking the appropriate measure based upon study design, resources and the sleep disorder under study. We recommend each clinical trial in sleep medicine should include objective and self-report measurements.
- In the future, sleep medicine assessment techniques are likely to be more portable and less expensive, potentially allowing long-term measurement/monitoring. These emerging technologies will have to be compared with existing gold standards.

References

- 1 American Academy of sleep medicine. *International Classification of Sleep Disorders: Diagnostic and Coding Manual (2nd edition)*. American Academy of Sleep Medicine, Westchester, IL, USA, 23 (2005).
- 2 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (4th edition)*. American Psychiatric Publishing, Inc., Arlington, VA, USA, 992 (2000).
- 3 Reynolds CF 3rd, Redline S, DSM-V Sleep-Wake Disorders Workgroup and Advisors. The DSM-V sleep-wake disorders nosology: an update and an invitation to the sleep community. *Sleep* 33(1), 10–11 (2010).
- 4 Ohayon MM, Dauvilliers Y, Reynolds CF 3rd. Operational definitions and algorithms for excessive sleepiness in the general population: implications for DSM-5 nosology. *Arch. Gen. Psychiatry* 69(1), 71–79 (2012).
- 5 Epstein LJ, Kristo D, Strollo PJ Jr *et al.* Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J. Clin. Sleep Med.* 5(3), 263–276 (2009).
- 6 Zeitzer JM, Nishino S, Mignot E. The neurobiology of hypocretins (orexins), narcolepsy and related therapeutic interventions. *Trends Pharmacol. Sci.* 27(7), 368–374 (2006).
- 7 Phillips BA, Schmitt FA, Berry DT, Lamb DG, Amin M, Cook YR. Treatment of obstructive sleep apnea. A preliminary report comparing nasal CPAP to nasal oxygen in patients with mild OSA. *Chest* 98(2), 325–330 (1990).
- 8 Carskadon MA, Dement WC. Sleep tendency: an objective measure of sleep loss. *Sleep Res.* 6, 200 (1977).
- 9 Mitler MM, van den Hoed J, Carskadon MA *et al.* REM sleep episodes during the Multiple Sleep Latency Test in narcoleptic patients. *Electroencephalogr. Clin. Neurophysiol.* 46(4), 479–481 (1979).
- 10 Littner MR, Kushida C, Wise M *et al.* Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. *Sleep* 28(1), 113–121 (2005).
- 11 Dauvilliers Y, Gosselin A, Paquet J, Touchon J, Billiard M, Montplaisir J. Effect of age on MSLT results in patients with narcolepsy-cataplexy. *Neurology* 62(1), 46–50 (2004).
- 12 Nardone R, Bergmann J, Lochner P *et al.* Modafinil reverses hypoexcitability of the motor cortex in narcoleptic patients: a TMS study. *Sleep Medicine* 11(9), 870–875 (2010).
- 13 Mitler MM, Gujavarty KS, Browman CP. Maintenance of wakefulness test: a polysomnographic technique for evaluation treatment efficacy in patients with excessive somnolence. *Electroencephalogr. Clin. Neurophysiol.* 53(6), 658–661 (1982).
- 14 Doghramji K, Mitler MM, Sangal RB *et al.* A normative study of the maintenance of wakefulness test (MWT). *Electroencephalogr. Clin. Neurophysiol.* 103(5), 554–562 (1997).
- 15 Black J, Pardi D, Hornfeldt CS, Inhaber N. The nightly administration of sodium oxybate results in significant reduction in the nocturnal sleep disruption of patients with narcolepsy. *Sleep Med.* 10(8), 829–835 (2009).
- 16 Black J, Houghton WC. Sodium oxybate improves excessive daytime sleepiness in narcolepsy. *Sleep* 29(7), 939–946 (2006).
- 17 Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep* 14(6), 540–545 (1991).
- 18 Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep* 15(4), 376–381 (1992).
- 19 Vana KD, Silva GE, Goldberg R. Predictive abilities of the STOP-Bang and Epworth Sleepiness Scale in identifying sleep clinic patients at high risk for obstructive sleep apnea. *Res. Nurs. Health* 36(1), 84–94 (2013).
- 20 Doshi R, Westbrook P. Nasal expiratory positive airway pressure (EPAP) for the treatment of obstructive sleep apnea: a review of clinical studies of provent therapy. *Respir. Ther.* 6(4), 45–49 (2011).
- 21 Berry RB, Kryger MH, Massie CA. A novel nasal expiratory positive airway pressure (EPAP) device for the treatment of obstructive sleep apnea: a randomized controlled trial. *Sleep* 34(4), 479–485 (2011).
- 22 Kryger MH, Berry RB, Massie CA. Long-term use of a nasal expiratory positive airway pressure (EPAP) device as a treatment for obstructive sleep apnea (OSA). *J. Clin. Sleep. Med.* 7(5), 449B–453B (2011).
- 23 Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. *Psychophysiology* 10(4), 431–436 (1973).
- 24 Maclean AW, Fekken GC, Saskin P, Knowles JB. Psychometric evaluation of the Stanford Sleepiness Scale. *J. Sleep Res.* 1(1), 35–39 (1992).
- 25 Shen J, Hossain N, Streiner DL *et al.* Excessive daytime sleepiness and fatigue in depressed patients and therapeutic response of a sedating antidepressant. *J. Affect Disord.* 134(1–3), 421–426 (2011).
- 26 Weaver TE, Laizner AM, Evans LK *et al.* An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep* 20(10), 835–843 (1997).
- 27 Chasens ER, Ratcliffe SJ, Weaver TE. Development of the FOSQ-10: a short version of the Functional Outcomes of Sleep Questionnaire. *Sleep* 32(7), 915–919 (2009).
- 28 Weaver TE, Mancini C, Maislin G *et al.* Continuous positive airway pressure treatment of sleepy patients with milder obstructive sleep apnea: results of the CPAP Apnea Trial North American Program (CATNAP) randomized clinical trial. *Am. J. Respir. Crit. Care Med.* 186(7), 677–683 (2012).
- 29 Becker PM. Insomnia: prevalence, impact, pathogenesis, differential diagnosis, and evaluation. *Psychiatr. Clin. North Am.* 29(4), 855–870, Abstract VII (2006).
- 30 Edinger JD, Bonnet MH, Bootzin RR *et al.* Derivation of research diagnostic criteria for insomnia: report of an American Academy of Sleep Medicine Work Group. *Sleep* 27(8), 1567–1596 (2004).
- 31 Schutte-Rodin S, Broch L, Buysse D, Dorsey C, Sateia M. Clinical guideline for the evaluation and management of chronic insomnia in adults. *J. Clin. Sleep Med.* 4(5), 487–504 (2008).
- 32 Erman MK, Zammit G, Rubens R *et al.* A polysomnographic placebo-controlled evaluation of the efficacy and safety of eszopiclone relative to placebo and zolpidem in the treatment of primary insomnia. *J. Clin. Sleep Med.* 4(3), 229–234 (2008).
- 33 Kupfer DJ, Detre TP, Foster G, Tucker GJ, Delgado J. The application of Delgado's telemetric mobility recorder for human studies. *Behav. Biol.* 7(4), 585–590 (1972).
- 34 Tryon WW. Issues of validity in actigraphic sleep assessment. *Sleep* 27(1), 158–165.
- 35 Vallieres A, Morin CM. Actigraphy in the assessment of insomnia. *Sleep* 26(7), 902–906 (2003).
- 36 Morgenthaler T, Alessi C, Friedman L *et al.* Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. *Sleep* 30(4), 519–529 (2007).
- 37 Natale V, Plazzi G, Martoni M. Actigraphy in the assessment of insomnia: a quantitative approach. *Sleep* 32(6), 767–771 (2009).

- 38 Weiss AR, Johnson NL, Berger NA, Redline S. Validity of activity-based devices to estimate sleep. *J. Clin. Sleep Med.* 6(4), 336–342 (2010).
- 39 Pollak CP, Tryon WW, Nagaraja H, Dzwonczyk R. How accurately does wrist actigraphy identify the states of sleep and wakefulness? *Sleep* 24(8), 957–965 (2001).
- 40 Gross CR, Kreitzer MJ, Reilly-Spong M *et al.* Mindfulness-Based Stress Reduction Versus Pharmacotherapy for Chronic Primary Insomnia: a randomized controlled clinical trial. *Explore (NY)* 7(2), 76–87 (2011).
- 41 Carney CE, Buysse DJ, Ancoli-Israel S *et al.* The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep* 35(2), 287–302 (2012).
- 42 McCall C, McCall WV. Comparison of actigraphy with polysomnography and sleep logs in depressed insomniacs. *J. Sleep Res.* 21(1), 122–127 (2012).
- 43 McCall C, McCall WV. Objective vs. subjective measurements of sleep in depressed insomniacs: first night effect or reverse first night effect? *J. Clin. Sleep Med.* 8(1), 59–65 (2012).
- 44 Ancoli-Israel S, Krystal AD, McCall WV *et al.* A 12-week, randomized, double-blind, placebo-controlled study evaluating the effect of eszopiclone 2 mg on sleep/wake function in older adults with primary and comorbid insomnia. *Sleep* 33(2), 225–234 (2010).
- 45 Bastien CH, Vallieres A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med.* 2(4), 297–307 (2001).
- 46 Morin CM, Belleville G, Belanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep* 34(5), 601–608 (2011).
- 47 Swift N, Stewart R, Andiappan M, Smith A, Espie CA, Brown JSL. The effectiveness of community day-long CBT-I workshops for participants with insomnia symptoms: a randomised controlled trial. *J. Sleep Res.* 21(3), 270–280 (2012).
- 48 Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 28(2), 193–213 (1989).
- 49 Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *J. Psychosom. Res.* 53(3), 737–740 (2002).
- 50 Morgan K, Gregory P, Tomeny M, David BM, Gascoigne C. Self-help treatment for insomnia symptoms associated with chronic conditions in older adults: a randomized controlled trial. *J. Am. Geriatr. Soc.* 60(10), 1803–1810 (2012).
- 51 Morin CM, Vallieres A, Ivers H. Dysfunctional beliefs and attitudes about sleep (DBAS): validation of a brief version (DBAS-16). *Sleep* 30(11), 1547–1554 (2007).
- 52 Carney CE, Edinger JD, Morin CM *et al.* Examining maladaptive beliefs about sleep across insomnia patient groups. *J. Psychosom. Res.* 68(1), 57–65 (2010).
- 53 Morin CM, Blais F, Savard J. Are changes in beliefs and attitudes about sleep related to sleep improvements in the treatment of insomnia? *Behav. Res. Ther.* 40(7), 741–752 (2002).
- 54 Edinger JD, Carney CE, Wohlgenuth WK. Pretherapy cognitive dispositions and treatment outcome in cognitive behavior therapy for insomnia. *Behav. Ther.* 39(4), 406–416 (2008).
- 55 Dolan DC, Taylor DJ, Bramoweth AD, Rosenthal LD. Cognitive-behavioral therapy of insomnia: a clinical case series study of patients with co-morbid disorders and using hypnotic medications. *Behav. Res. Ther.* 48(4), 321–327 (2010).
- 56 Yang CM, Liao YS, Lin CM, Chou SL, Wang EN. Psychological and behavioral factors in patients with comorbid obstructive sleep apnea and insomnia. *J. Psychosom. Res.* 70(4), 355–361 (2011).
- 57 Sato M, Yamadera W, Matsushima M, Itoh H, Nakayama K. Clinical efficacy of individual cognitive behavior therapy for psychophysiological insomnia in 20 outpatients. *Psychiatry Clin. Neurosci.* 64(2), 187–195 (2010).
- 58 McCall WV, Batson N, Webster M *et al.* Nightmares and dysfunctional beliefs about sleep mediate the effect of insomnia symptoms on suicidal ideation. *J. Clin. Sleep Med.* 9(2), 135–140 (2013).
- 59 Morgenthaler TI, Lee-Chiong T, Alessi C *et al.* Practice parameters for the clinical evaluation and treatment of circadian rhythm sleep disorders. An American Academy of Sleep Medicine report. *Sleep* 30(11), 1445–1459 (2007).
- 60 Gradisar M, Dohnt H, Gardner G *et al.* A randomized controlled trial of cognitive-behavior therapy plus bright light therapy for adolescent delayed sleep phase disorder. *Sleep* 34(12), 1671–1680 (2011).
- 61 Crowley SJ, Eastman CI. Melatonin in the afternoons of a gradually advancing sleep schedule enhances the circadian rhythm phase advance. *Psychopharmacology* 225(4), 825–837 (2013).
- 62 Cuthill FM, Espie CA. Sensitivity and specificity of procedures for the differential diagnosis of epileptic and non-epileptic seizures: a systematic review. *Seizure* 14(5), 293–303 (2005).
- 63 Krakow B. Nightmare complaints in treatment-seeking patients in clinical sleep medicine settings: diagnostic and treatment implications. *Sleep* 29(10), 1313–1319 (2006).