

A Systematic Review of Instruments for the Assessment of Insomnia in Adults

This article was published in the following Dove Press journal:
Nature and Science of Sleep

Raja Mahamade Ali

Monica Zolezzi 

Ahmed Awaisu 

College of Pharmacy, QU Health, Qatar
University, Doha 2713, Qatar

Introduction: Self-reported sleep instruments remain the most practical methods for the assessment of insomnia in clinical practice. This systematic review aims to identify, describe and summarize the psychometric properties of questionnaires available for the assessment of insomnia in the adult population. In addition, the review also aimed to identify sleep instruments available in the Arabic language.

Methods: A systematic literature search was conducted using the following electronic databases: PubMed, EMBASE, ProQuest Central, SCOPUS, and Google Scholar. The quality assessment of the instruments was conducted using two established international criteria.

Results: One hundred and seven articles were selected for inclusion, from which 31 instruments were identified and categorized based on the constructs they assess as: (1) screening for insomnia (n=14); (2) measuring the consequences of insomnia (n=8); (3) assessing the cognitive aspects of insomnia (n= 5); and (4) assessing sleep hygiene (n= 4). The review of the psychometric properties showed that the Insomnia Severity Index and the Functional Outcomes of Sleep Questionnaire were the most extensively evaluated instrument. Criterion validity and reliability measures were the most commonly reported properties. Only four of the identified instruments were available in Arabic.

Discussion: Overall, the findings of this study indicate ample availability of sleep instruments. However, psychometric testing for several of the available sleep instruments remains incomplete, particularly responsiveness and interpretability. Our findings suggest that future studies should focus on reporting more psychometric measures to ensure the trustworthiness of these instruments.

Keywords: insomnia, sleep hygiene, sleep quality, questionnaires, psychometric properties

Introduction

Worldwide reports suggest that around one-third of the adult population complains of insomnia symptoms.¹⁻² Insomnia is characterized by persistent sleep difficulty despite adequate sleep opportunity and associated daytime dysfunction.³ Insomnia has been associated with increased rates of road accidents, lower productivity and work absenteeism.^{4,5} Several factors are known to contribute to the development or worsening of insomnia symptoms. Current findings from the literature suggest that negative cognitive processes including worry, rumination and catastrophizing thoughts are associated with worse sleep quality and insomnia.⁶ These processes have been shown to fuel anxiety and arousal resulting in delayed sleep onset and shorter sleep duration.⁷ Similarly, poor sleep hygiene, which consists of a combination of behavioral practices and environmental conditions which improves sleep, is common among insomniacs.⁸⁻¹¹

Correspondence: Monica Zolezzi
College of Pharmacy, QU Health Qatar
University, PO Box 2713, Doha, Qatar
Tel +974 4403 5623
Fax +974 4403 5551
Email mzolezzi@qu.edu.qa

Studies suggest that chronic insomnia is an independent risk factor for developing mental illnesses in otherwise healthy individuals.^{12,13} Chronic insomnia has also been reported to increase the probability of developing chronic medical conditions, such as hypertension, cardiovascular disease and type 2 diabetes.^{14,15} Therefore, early detection and management of insomnia is important to minimize these associated risks. Although polysomnography (PSG) is considered the gold standard method for evaluating insomnia, it is not routinely used as it requires a specialized setting and equipment, and it is often labour intensive. Wrist actigraphy is another objective tool for the assessment of insomnia, but is limited by its lack of specificity.^{16,17} Consequently, self-reported sleep instruments remain the most practical methods for the assessment of insomnia in clinical practice.^{16,18} There are several systematic reviews on a range of sleep instruments utilized for the assessment of sleep dysfunction in a variety of medical and neuropsychiatric disorders, or in specialized populations.^{19–21} In these reviews, there was insufficient evaluation of the psychometric properties of the instruments that were identified, and none of them included instruments which assess sleep hygiene practices in the adult population. Therefore, the main objective of this study is to present an updated systematic review of the literature on validated self-reported instruments used for the assessment of different dimensions of sleep in the adult population. In addition, the review also aimed to identify sleep instruments available in Arabic language.

Methods

Search methods

Data Sources and Search Strategy

A comprehensive systematic search was conducted to identify studies reporting the development and/or validation of instruments for the assessment of self-reported sleep and sleep hygiene in the adult population. The five databases and search engines utilized included PubMed (1966 - April 2018), EMBASE (1980 - April 2018), ProQuest Central (1947–2018), SCOPUS (1966 - April 2018), and Google Scholar (till April 2018). Grey literature was also searched by reviewing conference proceedings and abstracts of the Canadian Sleep Society and the American Academy of Sleep Medicine published in the period from January 2014 to December 2017. Additionally, a hand search of the bibliographies of the articles identified through the electronic databases search

was undertaken. Search terms used in the electronic databases search were classified into three categories related to: sleep dysfunction (Category A), the instrument for assessment (Category B), and validation and psychometric properties (Category C). Terms from Category C were not used in databases such as PubMed which offered “validation studies” as one of the limits or filters. Publication language was limited to English and no limits were imposed on the publication year. Examples of full search strategies for two of the electronic databases can be found under the [supplementary material](#). In addition to the English search, two separate searches were conducted to identify sleep instruments developed in Arabic language or translated into Arabic. The first used the same search terms as those used in the original search, in addition to the word “Arabic”. The second was conducted in Arabic language in the following databases: PubMed, Google Scholar and Dar Al-mandumah using the same search terms as those used in the original search.

Inclusion Criteria

Articles selected for full review were those providing results of validation studies or reporting psychometric properties of instruments and questionnaires assessing characteristics of sleep (quality, quantity, nocturnal awakenings), daytime consequences of poor sleep, or sleep hygiene in the adult population. The inclusion criteria were limited to instruments completed by self-report and to instruments written in English or Arabic.

Exclusion Criteria

Studies were excluded from the review if they included evaluating instruments designed to measure sleep disorders other than insomnia (eg obstructive sleep apnea, restless leg syndrome, etc.), those using sleep items as subdomains of an instrument assessing a condition other than insomnia, studies which focused on pediatric/geriatric populations, studies describing instruments designed to be completed by clinician or caregiver and not by the patient, using instruments developed in languages other than English or Arabic, or describing instruments not psychometrically validated.

Data Collection and Analysis

Data Collection Process and Data Extraction

Duplicate citations were removed after obtaining the initial records of relevant articles from different databases. Titles and abstracts of the articles were then screened for

relevance. Full-text of eligible articles were obtained and screened against the inclusion and exclusion criteria. Data were extracted from the eligible studies according to eight key attributes, as established by the Scientific Advisory Committee, Medical Outcomes Trust (SAC-MOT).²² The extraction tool used in this study included the following elements: instrument's name, authors, conceptual framework (domains and purpose), psychometric properties, validation population, general description of the instrument (number of items, scale, scoring, response format, burden), and cultural and linguistic adaptation. The psychometric properties were extracted from studies describing the validation of the original (English or Arabic) version of the instruments. The psychometric properties extracted from the validation studies included validity, reliability, responsiveness, and interpretability. In addition, instruments' attributes related to the validation sample were extracted. Data extraction was done by one reviewer (RA) and reviewed by a second author (MZ). The psychometric properties for instruments validated in specific populations were not extracted.

Quality Assessment

A modified version of the criteria developed by Terwee et al²³ was used to assess the quality of the sleep instruments. The Terwee et al. criteria evaluates 8 psychometric properties including validity (content, construct and criterion), reliability (internal consistency and reproducibility), responsiveness, interpretability and floor and ceiling effects.²³ For the purposes of this study, the criteria for the floor and ceiling effects and the agreement component of the reproducibility were not used, primarily because, for the most part, these measures were not reported in the validation studies included in this review. In addition, Pearson's Correlation Coefficient was used for rating the reliability, responsiveness, criterion and construct validity.

Results

As illustrated in [Figure 1](#), a total of 4453 citations were retrieved from the search. One hundred and seven articles were deemed suitable for inclusion in the review. These articles included 31 distinct sleep instruments. Of the 107 articles included, 47 discussed the validation process of the instruments in English, while the remaining 60 articles reviewed the translation and cultural adaptation of these instruments into a variety of languages and populations. The two additional searches to identify sleep instruments in the Arabic language did not identify any additional

results than those derived from the original search. [Table 1](#) summarizes the results of the validation studies (n=47) related to the 31 sleep instruments included in this review. The table also describes the psychometric properties of the instruments and the characteristics of the populations in which they were validated. [Table 2](#) provides a detailed description of the characteristics of the 31 instruments.

As summarized in [Tables 1](#) and [2](#), the majority of the instruments identified contain less than 20 items. The longest of these instruments is the Sleep Practices and Attitudes Questionnaire (SPAQ), which contains 151 questions divided into 16 different domains.⁷⁵ While the Karolinska Sleepiness Scale (KSS),²⁴ the Minimal Insomnia Symptoms Scale (MISS),^{25,26} and the Jenkins Sleep Scale (JSS)²⁷⁻⁷³ are the shortest instruments identified, consisting of one, three, and four questions, respectively. Some of the instruments have multiple versions, each consisting of a different number of questions [eg, the Functional Outcomes of Sleep Questionnaire (FOSQ)-30 and the FOSQ-10 have 30 and 10 questions, respectively].^{74,28} The majority of the instruments use Likert-type scales as response options to generate scores. The time needed to complete an instrument was not reported in the majority of the studies reviewed. However, the response burden, wherever reported, did not exceed 10 minutes. The recall period for the majority of the instruments was one month, except for the Insomnia Severity Index (ISI)^{29,30} and the Sleep Functional Impact Scale (SFIS)³¹ which had a recall period of 2 weeks and 1 week, respectively. Only four of the identified instruments were available in Arabic, three of which (PSQI,³² ISI,³³ and ESS³⁴) were originally developed in English, but translated and validated in Arabic-speaking populations. The Arabic Scale of Insomnia (ASI)³⁵ is the only instrument which was originally developed in Arabic.

The instruments were classified into four categories based on the outcomes that were assessed in the 107 studies, as follows: (1) instruments screening for insomnia symptoms (n=14); (2) instruments assessing consequences of poor sleep (n=8); (3) instruments assessing the cognitive aspect of insomnia (n=5) and; (4) instruments evaluating sleep hygiene (n=4). A detailed description of the reported psychometric properties for the four categories of sleep instruments retrieved from the studies included in this review is provided in [Table 3](#).

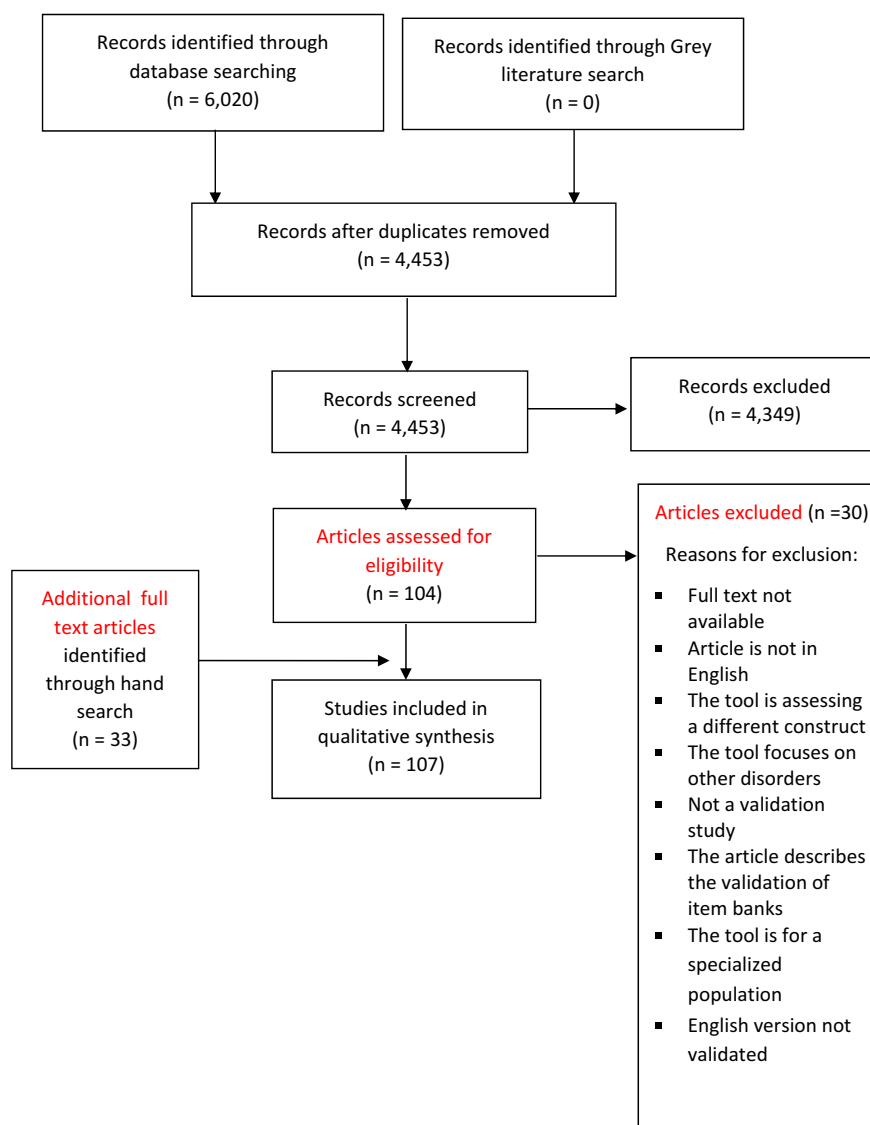


Figure 1 Flow chart of the systematic review process.

Instruments Screening for Insomnia Symptoms

From the 14 instruments screening for insomnia symptoms, two [the Women's Health Initiative Insomnia Rating Scale (WHIIRS)^{36,37} and the Restorative Sleep Questionnaire (RSQ)³⁸] were dimension-specific and focus on evaluating the sleep quality through assessing problems with initiation and maintenance of sleep. The remaining 12 instruments were multidimensional, assessing sleep quality and consequences of poor sleep. Some of the identified instruments also measure other specific dimensions such as satisfaction with sleep in the Bergen Insomnia Scale (BIS)³⁹ and sleep environment in the Insomnia Screening Scale (ISS).⁴⁰ While the Daily

Cognitive-Communication and Sleep Profile (DCCASP)⁴¹ measures the daily fluctuation in sleep quality and evaluates the impact of these fluctuations on the individual's cognitive functioning and communication, while the General Sleep Disturbance Scale (GSDS)^{42,43} assesses sleeping patterns and the use of sleep aids during the past month.

Instruments Assessing Consequences of Poor Sleep

These instruments measure an individual's functioning and daytime performance. Two of these instruments [the Epworth Sleepiness Scale (ESS)⁴⁴⁻⁴⁶ and the KSS)^{24,47} focus on assessing sleepiness only, whereas

Table 1 Studies Testing the Psychometric Properties of Extracted Sleep Instruments

Author, year	Instruments	Sample	Validity	Reliability	Responsiveness and Interpretability
Palleisen et al, 2008 ³⁹	The Bergen insomnia scale (BIS)	University students: (n=320) Community sample: (n=5,000) Patient sample: (n=225)	<i>Convergent and discriminant:</i> - BIS significantly associated with Athens Insomnia scale ($r=0.79$) and the PSQI ($r=0.73$). - BIS had lower correlation with Beck depression and anxiety inventories ($r=0.55$ and 0.32 respectively).	<i>Internal consistency:</i> Students: $\alpha = 0.79$ Community: $\alpha = 0.87$ Patient sample: $\alpha = 0.80$ -Test retest (after 2 weeks) Students (n=200): $r = 0.77$	NR
Yeh et al. 2012 ⁴⁰	The Insomnia Screening Scale (ISS)	Study 1: (n=162) -Primary insomniacs -Healthy participants Study 2: (n=262) Community sample (paediatrics, adults and elderly)	<i>Concurrent/criterion validity:</i> - ISS significantly correlated with Insomnia severity index (ISI) and PSQI (PSQI): ($r=0.87$ and $r=0.85$ respectively). - Daytime functioning domain significantly correlated with ISI and PSQI ($r=0.63$ and 0.68 respectively). - Sleep environment domain negatively correlated with PSQI and ISI: ($r = -0.31$ and -0.25 respectively). - Sleep opportunity negatively correlated with ISI and PSQI: ($r = -0.37$ and $r = -0.41$, respectively). <i>External validity:</i> Community sample: 13.36% (inadequate sleep opportunity and sleep environment), and 37.02% met the criteria of insomnia and daytime function.	<i>Internal consistency:</i> -Insomnia symptoms: $\alpha = 0.98$ -Daytime function: $\alpha = 0.94$ -Sleep environment: $\alpha = 0.90$ -Sleep opportunity: $\alpha = 0.87$	<i>Interpretability:</i> ISS demonstrated sensitivity and specificity of 0.89 and 0.59 respectively Cut-off scores: -Sum of sleep environment and opportunity subscales scores= 27 -Sum of insomnia symptoms and daytime functioning subscales scores=42
Kato T., 2013 ⁷⁰	Sleep Quality Questionnaire (SQQ)	Full-time employees and college students	<i>Content validity:</i> confirmed by two Japanese experts in stress research. <i>Convergent validity:</i> (n=370) SQQ subscales (Sleep difficulty and daytime sleepiness) with MOS sleep scale ($r_s = 0.37$ and 0.43 , $p < 0.001$). The Daytime Sleepiness subscale score with Epworth Sleepiness Scale (ESS) score ($r = 0.47$, $p < 0.001$). <i>Incremental validity:</i> (n=346) SQQ subscales with the GHQ-12, CES-D Scale, FSS, and SWLS scores significant ΔR scores (0.403, 0.313, 0.408 and 0.054 respectively) with $p < 0.001$.	<i>Internal consistency:</i> Daytime sleepiness subscale: -Student sample: $\alpha = 0.83$ -Employees sample: $\alpha = 0.84$ Sleep Difficulty subscale: -Students sample: $\alpha = 0.74$ -Employees sample: $\alpha = 0.77$ Test-retest (over 8 weeks) Daytime Sleepiness subscale: ($r = 0.76$) Sleep Difficulty subscale: ($r = 0.79$)	NR

(Continued)

Table 1 (Continued).

Author, year	Instruments	Sample	Validity	Reliability	Responsiveness and Interpretability
Levine et al, 2003 (study a) ³⁶ Levine et al, 2003 (study b) ³⁷	Women's Health Initiative Insomnia Rating Scale (WHIRS)	Postmenopausal women	Content analysis: The WHIRS items corresponded to most of insomnia characteristics noted in the nosologies (eg International Classification of Sleep Disorders) and the literature. Construct validity: - Small correlation between the CES-D and the WHIRS: $r = 0.29$ - The WHIRS mean in the largest CES-D category ($M = 10.3$ for Category 12) was 1.8 times that in the smallest depression category (ie, $M = 5.7$ for Category 0) - The SF-36 subscales were linearly correlated with the WHIRS: Cohen's f value was 0.273, ($p < 0.0001$). - For night sweats and hot flushes with WHIRS, the values of Cohen's f were, respectively, 0.205 and 0.157.	Internal consistency: $\alpha = 0.786$. • 89.3% of the samples had reliability coefficients ≥ 0.75 Test-retest: Same day administration ($r = 0.96$) Tests after >1 year ($r = 0.66$)	NR
Drake et al, 2014 ³⁸	Restorative sleep questionnaire (RSQ)	Community based sample Patients with primary insomnia Non-refreshed sleepers (NRS)	Content validity: - Key concepts developed through patient focus groups and patient interviews. Concepts reviewed by two expert panels. Convergent/divergent: - RSQ-Daily (RSQ-D) with Leeds Sleep Evaluation Questionnaire scores: positive and significant correlation with all domains ($r \geq 0.40$, $p \leq 0.006$) except for Getting to Sleep scores ($r = 0.27$; $p = 0.079$). - RSQ-D with Subjective Sleep questionnaire: Sleep quality ($r = 0.59$; $p < 0.001$); TST ($r = 0.32$; $p = 0.036$) - RSQ-D with vitality questionnaire: ($r = 0.61$; $p < 0.001$) - RSQ with PSG: Latency to persistent sleep ($r \geq 0.20$), TST and sleep efficiency ($r \geq 0.26$), WASO ($r \leq 0.020$) and total wake time ($r \leq 0.26$) significant p-value.	Internal consistency: RSQ-D: $\alpha = 0.91$ RSQ-W: $\alpha = 0.90$ Test-retest (responses on consecutive days measures) RSQ-D and RSQ-W: $r > 0.80$	NR

Buyse et al, 1989⁶⁷ Grandner et al, 2006⁶⁸ Backhaus et al, 2002⁶⁹	Pittsburgh sleep quality index (PSQI)	- Group 1: healthy control subjects (n=52) - Group 2: patients with major depressive disorder (n=34). - Group 3: Clinical sample physician-referred outpatients at the Sleep Evaluation Center (n = 45) - Non-clinical sample (younger and older adults) - Primary insomniacs - Healthy controls	<i>Discriminating between patients and controls:</i> - Global PSQI scores differed significantly between subject groups and control group subjects differed from all patient groups. - Patients with DIMS had significantly higher scores than patients with DOES patients. A significant difference in PSQI components' scores were found between the control group and both DIMS and depressed groups. PSQI scores also differed on 3 components (sleep disturbances, daytime dysfunction, and sleep quality) with DOES patients. - A significant difference in all PSQI component scores except sleep disturbance were identified between DOES and DIMS patients. While DOES and depressed patients differed on all components' scores except sleep disturbance and daytime dysfunction. <i>Diagnostic validity:</i> (Buyse et al, 1989) A cutoff score of 5 correctly determined the sleep quality for 88.5 of all participants ($\kappa = 0.75$, $p < 0.001$) reporting a sensitivity of 89.6% and a specificity of 86.5%. (Backhaus et al, 2002) - PSQI global score > 5 resulted in a sensitivity of 98.7% and specificity of 84.4%. <i>Criterion/concurrent validity:</i> (Buyse et al, 1989) PSQI scores and PSG: - Significant positive correlation between PSG and PSQI only for sleep latency ($r = 0.33$, $p < 0.001$) PSQI global score and PSG: - Objective sleep latency ($r = 0.20$, $p < 0.01$), weak correlation - The global PSQI score correlated only with REM% in controls ($r = 0.34$, $p < 0.006$) and number of arousals in depressives ($r = 0.47$, $p < 0.002$). (Backhaus et al, 2002) - PSQI was highly correlated with sleep log data ($r = 0.81$, $p < 0.001$ for sleep duration and $r = 0.71$, $p = 0.000$ for sleep onset latency). - The correlations between PSQI and PSG data was significant and lower ranging between $r = (-0.32$ to -0.33). <i>Construct validity:</i> - PSQI scores did not correlate significantly with actigraphic measures of sleep ($r \leq 0.13$). PSQI with sleep diary: (significant correlation) - Sleep efficiency ($r = -0.562$, $p < 0.01$), TST ($r = -0.307$, $p < 0.01$), WASO: ($r = 0.262$, $p < 0.01$). - Sleep latency: $r = 0.480$, $p < 0.01$ - CES-D: $r = 0.305$, $p < 0.01$	<i>Internal consistency:</i> Overall: $\alpha = 0.83$ (Buyse et al, 1989) Overall: $\alpha = 0.85$ (Backhaus et al, 2002) Test-retest: (Buyse et al, 1989) - Test retest (28 days apart): The correlation between the PSQI scores was $r = 0.85$ ($p < 0.001$). - Component scores: $r = 0.84-0.65$ ($p < 0.001$ for each component score). Test retest: (Backhaus et al, 2002) Overall: ($r = 0.87$, $p < 0.001$) - Short interval (2 days): $r = 0.90$, $p = 0.00$ - Longer interval (45.6 $\pm$ 18 days): $r = 0.86$, $p < 0.001$	<i>Interpretability:</i> Can distinguish between "good" and "poor" sleepers. The cut-off point is 5.
---	--	--	---	---	--

(Continued)

Table 1 (Continued).

Author, year	Instruments	Sample	Validity	Reliability	Responsiveness and Interpretability
Bastien et al, 2001 ²⁹ Morin et al, 2011 ³⁰	Insomnia severity index (ISI)	Study 1: Patients with insomnia complaint (n=145). Study 2: (n=78) Insomnia patient involved in a study assessing the efficacy of cognitive behavioral therapy. Morin et al, 2011 - Community sample (n= 959) - Clinical sample (n=183)	Criterion/concurrent validity: - ISI items with sleep diary variables: -Sleep onset latency ($r = 0.38$), WASO ($r = 0.35$) and Early morning awakening ($r = 0.35$) -Total ISI score and the sleep efficiency variable ($r = -0.19$) - ISI and sleep diary components $r = 0.32-0.55$ at baseline and $r = 0.50$ to 0.91 post-therapy (all p values < 0.05). - ISI variables with PSG variable: - Correlation ranged between $r = 0.07$ to 0.45 at pretreatment, and from 0.23 to 0.45 at post-treatment. Only the correlation for SOL was significant at pretreatment, whereas all correlations, but one (EMA) were significant at post-treatment ($p < 0.05$). - The correlations between the patient's and the clinician's versions of the ISI at the two assessment periods were all significant (p values < 0.01). Furthermore, the correlations between the patient's and the significant other versions of the ISI were also significant at the two assessment periods (p values < 0.01). Predictive validity: The clinician's ratings predicted best the patient's ISI total score at baseline, while at post-treatment, both the clinician and the sleep diary data were reliable predictors of the patient's total ISI score. $R^2 = 0.37$ ($p < 0.05$) at pretreatment, clinician: ($\beta = 0.52$) $R^2 = 0.61$ ($p < 0.05$) at post-treatment, clinician: ($\beta = -0.52$) Sleep diary: ($\beta = -0.34$) Content validity: A principal component analysis, using varimax rotation, explored the ISI content validity and the extent to which its components corresponded to insomnia's diagnostic criteria. Diagnostic validity: Morin et al, 2011 - <i>Subthreshold insomnia:</i> - A cutoff score of 8: - Sensitivity of 95.8% and 99.4% in the Community and Clinical samples respectively, and with specificity of 78.3% and 91.8%. - <i>Moderate to severe insomnia:</i> - A cutoff score of 15: - Specificity of 98.3% and 100% in the Community and Clinical samples, respectively, with a sensitivity of 47.7% and 78.1% respectively. - Community sample: cut-point of 10 (86.1% sensitivity and 87.7% specificity). - Clinical sample: cut-point of 11 was associated with 97.2% sensitivity and a perfect 100% specificity. Construct validity (convergent): - Insomnia severity on the ISI was positively correlated with the corresponding diary variable. - The ISI total score was significantly correlated with the PSQI total score, $r = 0.80$, $p < 0.05$. - Significant relationships were found with measures of anxiety and depression, different dimensions of fatigue. Significant correlations were also discovered with the SF-12, with a stronger association identified with the Mental components of the measure than with the Physical Health component.	Study 1: Internal consistency: $\alpha = 0.74$. Study 2: Internal consistency: The internal reliability coefficients did not change significantly from baseline to follow up (0.76 to 0.78, respectively). Morin et al, 2011 Community (Cronbach $\alpha = 0.90$) Clinical samples (Cronbach $\alpha = 0.91$).	Responsiveness: A significant reduction in ISI score was detected post-treatment in patients' (8.9 vs 15.4) and clinician's versions (7.7 vs 17.7). ISI sensitivity to detect clinical improvement (n=146) - Moderate improvement in insomnia was identified as a reduction in ISI score by > 7 points (60% sensitivity, 70% specificity). - Marked/higher improvement in sleep was identified as a decrease in ISI score by > 8 points (64% sensitivity, 80% specificity). Interpretability: The scores of the ISI represent the following: 0-7: lack of insomnia, 8-14: Subthreshold insomnia 15-21: Moderate insomnia 22-28: Severe insomnia. Morin et al (2011) suggested clinically significant insomnia could be identified by ISI scores > 10 in community samples and scores > 11 in clinical settings.

Jenkins et al, 1988 ²⁷ Nassermoaddeli et al, 2005 ⁷³ Jerlock et al, 2006 ⁷²	The Jenkins Sleep Scale (JSS)	-Air traffic controllers (n=250 men; 25-49 years old) -Patients recovering after from cardiac surgery (n=500 patients).	Construct validity: Spielberger's state anxiety: $r = 0.37$ POMS-Depression: $r = 0.35$ POMS-Hostility: $r = 0.29$ POMS-Vigor: $r = -0.24$ POMS-Fatigue: $r = 0.46$ Positive well-being: $r = -0.22$	Internal consistency: $\alpha = 0.79$ Test-retest reliability: $r = 0.59$ Nassermoaddeli et al, 2005 Internal consistency: $\alpha = 0.77$ Jerlock et al, 2006 Internal consistency: $\alpha = 0.80$	NR
Soldatos et al, 2000 ⁶⁵ Soldatos et al, 2001 ⁶⁵	Athens Insomnia Scale (AIS)	299 subjects Consisting of: 1-Primary insomniacs 2- Psychiatric patients (both inpatients and outpatients) 3-Healthy subjects	External validity: Sleep problems scale with: AIS-8: $r = 0.90$, $p < 0.001$ AIS-5: $r = 0.85$, $p < 0.001$ Soldatos et al, 2001 Diagnostic validity: - AIS in general population, for a Cut-off score of 6 only 1% of those responders with insomnia are misdiagnosed (ie NPV= 99%) and (PPV=41%). -Psychiatric population: 8% (NPV=92%) and (PPV=86%). -A total score of 6 or higher in the AIS was shown to correctly identify 90% of the study subjects as suffering or not from "nonorganic insomnia" according to ICD-10. -the AIS in overall case identification corresponded to high indices of both sensitivity (93%) and specificity (85%).	Internal consistency: AIS-8: $\alpha = 0.89$, AIS-5: $\alpha = 0.87$ Test retest (n=194; within one week): AIS-8: $r = 0.89$, $p < 0.001$ AIS-5: $r = 0.88$, $p < 0.001$	Interpretability: Cut-off score of 6 General population (NPV= 99%, PPV=41%) Psychiatric population (NPV=92%, PPV=86%) with sensitivity (93%) and specificity (85%). -A total score of 6 or higher in the AIS was shown to correctly identify 90% of the subjects' sleep quality.
Broman et al, 2008 ²⁵ Westergren et al, 2015 ²⁶	Minimal Insomnia Symptom Scale (MISS)	Subjects selected randomly (n=1379) Age range: 20-64 (Sweden).	Criterion validity: The correlation between BNSQ question about sleep quality and MISS total score was high with $r = 0.76$. Westergren et al, 2015 The difference between the adult and elderly samples was lower for the originally recommended ≥ 6 points cut-off (0.09 logits).	Internal consistency: Total: $\alpha = 0.73$ Test retest reliability: (6 months apart) High correlation was found with an ICC of 0.79.	Responsiveness: Sensitivity to change Paired t tests revealed that there was a strong trend for increase in score in subjects who deteriorated ($m = +0.80$; $t = 2.0$; $p = 0.053$). There was also a significant decrease in MISS scores among subjects who improved ($m = -1.06$; $t = 2.9$; $p < 0.01$). Interpretability: A cut-off score of ≥ 6 on the MISS identify insomniacs in the general adult population (sensitivity 0.82; specificity 0.86), PPV 0.44; NPV 0.97.

(Continued)

Table 1 (Continued).

Author, year	Instruments	Sample	Validity	Reliability	Responsiveness and Interpretability
Esple et al, 2014 ⁶⁴	Sleep Condition Indicator (SCI)	Samples from 5 validation studies: - The Great British Sleep Survey (GBSS): adults > 18 - The GBSS+ - TV sample - Glasgow Science Centre data (n=256) - A randomized controlled trial (RCT) sample (n=164) recruited into a placebo-controlled evaluation of CBT	Content validation: - The DSM-5 was used to develop the questionnaire, consultations were conducted, and a draft was published on the American psychiatric association website. Concurrent/criterion validity: - SCI was negatively associated with the score of sleep quality questionnaires including PSQI ($r = -0.734$) and the ISI ($r = -0.793$). - Sleep condition was significantly associated with physical and mental health ($r = 0.222$ and $r = 0.335$ respectively). - Using the HADS scale, SCI was negatively correlated with symptoms of depression ($r = -0.426$) and anxiety ($r = -0.400$). But was higher than the associated detected in the RCT sample study (depression ($r = -0.267$), anxiety ($r = -0.236$) and stress ($r = -0.263$)). Concurrent and diagnostic validity: - A cut-off score ≤ 16 was diagnostic for 89% of individuals who were identified as insomniacs on the ISI scale (scores of ≥ 15), with a capacity of correctly classifying 82% of non-insomniacs. - These findings support the concurrent validity for the SCI and confirming that a score of ≤ 16 on the SCI could identify insomniac patients.	Internal consistency: - Cronbach's $\alpha = 0.857$ (the GBSS (range of α-if-item-deleted $0.822 - 0.860$)). - GBSS+ sample ($\alpha = 0.865$). The mean corrected item-total correlation was moderate ($r = 0.620$).	Interpretability: - A cut-off ≤ 16, was able to identify 89% of patients who had insomnia (ISI scores of ≥ 15). - However, an SCI score of > 16 was able to exclude 82% of individuals without insomnia
Lee, 1992 ⁴¹ Lee, 2007 ⁴²	General Sleep Disturbance Scale (GSDS)	Female nurses in different shifts (permanent day shift, permanent night shift and rotating shifts) N=760 Parents of hospitalized infants in the intensive care unit n=44	Lee, 2007 Content validity: Evaluated by multidisciplinary reviewers. Criterion/concurrent validity: 1 -Sleep diary: For both the Chinese and English versions, higher GSDS scores were correlated with self-reported lower sleep quality in sleep logs ($r = 0.41$), higher morning fatigue levels (Chinese version: $r = 0.42$, $p = 0.05$; English version: $r = 0.56$, $p = 0.006$). 2- Wrist actigraphy: less sleep efficiency (Chinese version: $r = -0.26$; English version: $r = -0.42$). Criterion/predictive validity: The participants' GSDS mean scores were positively correlated with their morning fatigue levels (Chinese version, $r = 0.42$, $p = 0.05$; English version, $r = 0.56$, $p = 0.006$), supporting the predictive validity of the GSDS.	Lee, 1992 Internal consistency: -Overall: $\alpha = 0.88$ -Subscales of quality of sleep, daytime sleepiness, and use of sleep aids were 0.62, 0.79, and 0.82, respectively. Lee, 2007 Internal consistencies: - Overall: Chinese version ($\alpha = 0.81$) - English version ($\alpha = 0.85$)	NR

Fung et al, 2014 ⁴¹	Daily Cognitive Communication and Sleep Profile (DCCASP)	University students' university of Toronto (n=59).	Criterion validity: DCCASP and PSQI: - Adequate criterion validity for the Sleep Quality domain of the DCCASP was established by comparing it to the Sleep Quality domain of the PSQI, with $r_s = 0.398$ ($p < 0.001$). -Positive correlation between sleep quality and each of the DCCASP domains: r_s (0.38-0.55) ($p < 0.0001$).	Internal consistency: Cronbach's α ranged between 0.864 and 0.938 across the seven domains of the DCCASP -Test-retest (repeated after 2 weeks) Concordance Correlation Coefficient (CCC) of each domain of the DCCASP was moderate, ranging from $r = 0.548$ to 0.742.	NR
Abdel-khalek, 2008 ³⁵	The Arabic Scale of insomnia (ASI)	Students and employees	Content validity: The initial draft of the ASI was reviewed by PhD holding faculty members and master students. Construct: Convergent validity: AIS was significantly correlated with Arabic sleep disorders scale and Jenkins sleep scale a correlation range of ($r = 0.56-0.94$, $p < 0.001$).	Internal consistency: $\alpha = 0.84-0.87$ Test-retest reliability (1 week apart): $\alpha = 0.70-0.83$	NR
Morrone et al, 2017 ⁵²	Maugeri Sleep Quality and Distress Inventory (MasQuDI-17)	Outpatients evaluated for sleep disorders in sleep centers of Northern Italy	Convergent validity and discriminant: ($p < 0.001$ for all) $r^2 = 0.5$ with the PSQI $r^2 = 0.15$ with ESS $r^2 = 0.39$ with anxiety as measured by A-D schedule $r^2 = 0.52$ with depression as measured by A-D schedule -Discriminating capacity: The difference in mean score between the healthy group and patients (9.31 and 14.15 respectively) was statistically significant ($p < 0.001$). - Differences were consistent with the type of pathological condition present (normal subject vs OSAS or INS or BSD group of subjects $p < 0.001$).	Internal consistency: $\alpha = 0.896$	NR

(Continued)

Table 1 (Continued).

Author, year	Instruments	Sample	Validity	Reliability	Responsiveness and Interpretability
Johns M., 1991 ⁴⁴ Johns M., 1992 ⁴⁵ Johns M., 1994 ⁴⁶	Epworth Sleepiness scale (ESS)	-Control (n=30) -Individuals with sleep disorders (n=150) -Third year medical students (n=104) -OSA patients treated with CPAP -Patients who had MSLT (n=44) -Spouses of participants who completed ESS (n=50)	Discriminating capacity: -Significant differences in ESS scores between the seven diagnostic groups were detected ($F=50.00$; $df=6,173$; $p<0.0001$). -Scores for OSA, narcolepsy and idiopathic hypersomnia were significantly higher than for controls ($p<0.001$) or primary snorers ($p<0.001$). - The insomniacs had significantly lower scores ($p<0.01$) than all groups except controls. Criterion validity: ESS (total score) with MSLT: -Mean Sleep latency (SL): ($\rho=-0.42$, $p<0.01$). [Significant negative correlation] -When the correlation between mean SL with ESS individual item scores was assessed only 3 items were significantly correlated ($p<0.05$) which were: ● The likelihood of dosing off "when sitting, inactive in a public place" ($\rho=-0.44$), when "sitting quietly after lunch without alcohol" ($\rho=-0.34$) and when "in a car, stopped for a few minutes in the traffic" ($\rho=-0.41$). Predictive validity: -All eight item scores as predictors of the mean SL in the MSLT were statistically significant ($r=0.639$, $p=0.01$). - Items 3 ($r=-0.44$, $p<0.01$) and 8 ($r=-0.41$, $p=0.01$) were significant independent predictors.	Internal consistency: $\alpha=0.88$ (patients) $\alpha=0.73$ (students) Test-retest reliability: (5 months apart) -For 87 students: $r=0.822$ ($p<0.001$). -Patient-Spouse paired item correlation: (mean $\rho=0.57$, $p<0.001$). - The paired (patient-Spouse) total ESS scores correlation (high) ($\rho=0.74$, $n=50$, $p<0.001$).	Responsiveness: After 3 months of treatment with CPAP: Treatment with nasal CPAP in 54 individuals with OSAS was associated with a change in ESS scores by 7.0 ± 5.2 (SD) following therapy, which was statistically significant ($t=-9.59$, $df=53$, $p<0.001$).
Akerstedt et al, 1990 ⁴⁴ Reyner, et al, 1998 ⁴⁷	Karolinska sleepiness scale (KSS)	8 male subjects	Criterion validity: KSS with VAS: -The maximum score on the KSS scale corresponds with the verbal anchor "extremely sleepy, fighting sleep," while the minimum value corresponds with the rating "alert." - The association between subjective sleepiness and the EEG/EOG variables was significant: ($r=0.29-0.65$, $p<0.05$). -Significant differences in various levels of subjective sleepiness were identified for all variables, except theta activity- during the test conducted under ambulatory conditions. A significant difference between maximum and minimum sleepiness was detected (Wilcoxon, $z>2.20$, $p<0.05$). -No changes were identified in the EEG/EOG before level 7 was reached on the KSS scale. For the test session with closed eyes, a significant variation was identified only with slow rolling eye movement ($\chi^2=13.6$, $p<0.01$). - Increasing KSS levels were highly significantly correlated with an increased likelihood of falling asleep (Pearson's $r=0.78$; $d=1043$; $p=0.001$).	NR	NR

Weaver et al. 1997 ⁷⁴ Weaver et al. 2009 ²⁸	Functional outcomes of sleep questionnaire (FOSQ)	FOSQ-30: Sample 1 (n = 153) Healthy individuals presenting with sleep complaint Samples 2 (n = 24) and 3 (n = 51): Patients with OSA FOSQ-10: Sample 1: (n=155) Participants with moderate to severe OSA on CPAP Sample 2: (n = 51) CPAP-treated OSA patients Normal subjects	FOSQ-30: <i>Face validity:</i> Seven judges with expertise in the areas of functional status instrument development and sleep problems rated the clinical relevance of each item and the instrument to DOES. <i>Content validity:</i> Determined by the proportion of items receiving a rating of at least three or four across all judges. <i>Construct validity:</i> -Subscale-to-subscale correlations range: $r = 0.52-0.86$ -Subscale-to-global FOSQ score intercorrelations ranged from $r = 0.78-0.86$. <i>Concurrent/criterion validity:</i> -FOSQ-30 global score with SIP total score (n=24): ($r = -0.50$, $p \leq 0.05$) -FOSQ-30 global with SF-36 sub-scales (n=51): no significant correlation -FOSQ-30 global score with SF36 role emotional functioning subscale ($r = 0.46$, $p \leq 0.01$) -FOSQ activity level subscale significantly correlated with SF36 physical functioning subscale. - The FOSQ social outcome subscale was significantly correlated with the SF-36 social function subscale ($r = 0.36$, $p \leq 0.05$) and SF36 mental health subscale ($r = 0.38$, $p \leq 0.01$). <i>Discriminant validity:</i> -Mean of FOSQ global score discriminated between normal sleepers and those with sleeping problems: (68.05 ± 21.24 and 89.59 ± 8.64 respectively with $p = 0.0004$) (T157 = -5.88, $p = 0.0001$). FOSQ-10: <i>Criterion validity:</i> Global scores comparison: -Before treatment with CPAP the FOSQ-10 global score was robustly associated with the FOSQ-30 Total score ($r = 0.96$, $p < 0.0001$). -The correlation between the post-treatment total scores of FOSQ-10 and the FOSQ-30 was $r = 0.97$ ($p < 0.0001$). <i>Known group validity:</i> -Baseline total score on the FOSQ-10 for the OSA group (mean = 12.48 ± 3.23) statistically differed ($t = 8.65$, $p < 0.0001$) from the normal values (mean = 17.81 ± 3.10).	FOSQ-30: <i>Internal consistency</i> (n=153): -Total: $\alpha = 0.95$ -Subscales ($\alpha = 0.86$ to $\alpha = 0.91$) - Item to total correlation range: 0.35–0.73 <i>Test retest reliability</i> (n=32): (Within 1 week) - Global score: $r = 0.90$ - Individual subscales ranged from $r = 0.81$ to $r = 0.90$ FOSQ-10: <i>Internal consistency:</i> $\alpha = 0.87$	Responsiveness: Following CPAP treatment, both the FOSQ-30 and the FOSQ-10 detected a large clinically meaningful change in the total score ($p < 0.0001$).
David, 2008 ⁵⁰	Occupational impact of sleep questionnaire (OISQ)	Community sample of 86 participants (age: 25–50 years) ● 43 meeting DSM IV criteria for primary insomnia (26 women & 17 men) ● 43 controls	Criterion: - At each time point OISQ scores positively correlate with PSQI (mean $r = 0.59$, $p < 0.001$). - Sleep diary variable: Work assessment scores negatively correlated with mean TST ($r = -0.47$, $p < 0.001$) and mean SE ($r = -0.56$, $p < 0.001$), and significantly and positively with mean WASO ($r = 0.66$, $p < 0.001$). -Significant negative correlation with SF-36 subscales ranging from $r = -0.21$ to -0.62 , $p < 0.001$.	<i>Internal consistency:</i> $\alpha = 0.93$ <i>Test-retest:</i> A mean difference of 10.82 between groups was found resulting in a significant main group effect ($F = 12.52$, $p < 0.001$). A consistent decrease of 10% is seen in insomniacs compared to control.	NR

(Continued)

Table 1 (Continued).

Author, year	Instruments	Sample	Validity	Reliability	Responsiveness and Interpretability
Bell et al, 2011 ³¹	The Sleep Functional Impact Scale (SFIS)	Primary insomniacs Healthy volunteers	<i>Content validity:</i> Was assessed through face to face interviews with patients. <i>Convergent/divergent validity:</i> (p < 0.001) -SFIS with ISI composite score: (r = 0.82), SFIS with PSQI composite: (r = 0.78), SFIS with FOSQ composite: (r = -0.69), SFIS with ESS total: (r = 0.46), SFIS with MOS- sleep problems indices (I and II): (r = 0.74), SFIS with MOS breathing and Snoring subscales: (r = 0.26 and 0.27 respectively) and with WPAI-GH Subscales (r = 0.31). <i>Known group validity</i> (n=430): -Higher SFIS score in the insomnia group compared to control (t=19.36, p< 0.0001), effect size=2.05 -Lower SFIS scores for those reporting good sleep compared not reporting good sleep. (t = -12.18, p< 0.001) (effect size = -1.04). -Higher SFIS score for participants reporting worse sleep quality on PSQI (t = 17.06, p< 0.0001) (effect size = 2.33).	<i>Internal consistency:</i> All the sample: $\alpha = 0.97$ Insomniacs: $\alpha = 0.95$	NR
Esple et al, 2000 ⁵¹	Sleep Disturbance Questionnaire (SDQ)	Chronic insomniacs	The factors score of SDQ correlated with the SDQ total (0.31 to 0.88).	<i>Internal consistency:</i> $\alpha = 0.67$	NR
Regestein et al, 1993 ⁴⁹ Pavlova et al, 2001 ⁴⁷	Hyperarousal Scale (H-scale)	-Primary insomnia patients -Hypersomnia -Delayed sleep syndrome -Normal subjects -Patients with refractory insomnia	<i>Criterion validity:</i> -H-Scale with neuroticism scale and extroversion-introversion scale: no significant correlation -Higher EEG activity in insomniacs compared to normal subjects (p < 0.05). -Significant correlation between hyperarousal score with alpha and non-alpha EEG activity: (r = 0.38, p ≤ 0.03). <i>Discriminating capacity:</i> (Pavlova et al, 2001) - The insomnia group had a mean Hyperarousal total score significantly higher than the normal group (F = 20.7, p < 0.001).	NR	NR

Espe et al, 2000⁵¹ Edinger et al, 2001⁵⁸ Chung et al, 2016⁵⁹	Dysfunctional Beliefs and Attitudes about Sleep Scale (DBAS)	Chronic insomniacs Normal subjects Insomniacs Individuals with insomnia derived for a randomized controlled trial of self-help CBT-I.	Construct validity: - DBAS-10 with DBAS-30: good correlation ($r = 0.826$). - Total score: DBAS-10 and SDQ ($r = 0.28$) Construct/convergent validity: Edinger et al, 2001 -DBAS with DBAS-10: Normal sleepers: $r = 0.84$, Insomniacs: $r = 0.80$ Discriminative validity: Insomniac subjects had significantly higher scores on the full-length DBAS scale ($F(1, 136) = 28.2$, $p < 0.0001$) with mean of (mean DBAS 35.5, SD 10.5) as compared to normal sleepers (mean DBAS 26.4, SD 9.7). - DBAS-10 scores of our insomnia sufferers (mean DBAS-10 43.3, SD 15.0) were significantly ($F(1, 136) = 15.9$, $p < 0.0001$) higher than those shown by our normal sleepers (mean DBAS-10 = 33.5, SD 14.0). Criterion/concurrent validity: Chung et al, 2016 - DBAS total score significantly correlated with several subscale scores of the ISI, PSQI total scores and HADS-anxiety and HADS-depression subscale scores. - The DBAS-10 "immediate consequences" subscale had significant correlation with ISI and HADS-anxiety subscale but was not related to HADS-depression subscale. The DBAS-30 and DBAS-16 "sleep expectation" subscales were only significantly correlated with ISI total score.	DBAS 30: Internal consistency: -Overall: $\alpha = 0.72$ -Only two subscales achieved significance: 1) "misattributions or amplifications of the consequences of insomnia"; $\alpha = 0.77$ 2) Subscale 4: "diminished perceptions of control and predictability of sleep." $\alpha = 0.41$. DBAS-10: -Internal consistency: $\alpha = 0.69$ -Internal consistency for factors I and II were 0.73 and 0.60. Edinger et al, 2001 Internal consistency: DBAS 30: Normal subjects: $\alpha = 0.81$ Insomniacs: $\alpha = 0.71$ DBAS-10: Normal subjects: $\alpha = 0.70$, Insomniacs: $\alpha = 0.53$ Chung et al, 2016 Internal consistency: DBAS-30, DBAS-16, and DBAS-10 with Cronbach α of 0.81, 0.80, and 0.73, respectively.	Responsiveness: -Changes in DBAS score following treatment were statistically comparable ($F(2,68) = 2.33$, $p < 0.10$) for all three treatment groups CBT, relaxation training (RT) or placebo control (PC). - DBAS-10 changes did differ statistically across the three treatment groups ($F(2, 68) = 4.69$, $p < 0.025$). -Post hoc comparisons showed that the CBT-treated insomniacs showed significantly greater decreases on the DBAS-10. - Significantly greater reduction in DBAS scores was found in participants allocated to CBT-I, except the DBAS-30 "attributions" subscale and DBAS-16 "medication" -Improvement in ISI scores by ≥ 8 points was associated with significant changes in DBAS total scores and on DBAS-30 and DBAS-10 subscale scores.
Kallestad et al, 2010⁵⁸	Insomnia daytime worry scale (IDWS)	Undergraduate students and senior high school students (Norway)	Predictive validity: The IDWS predicted insomnia severity over and above the other variables, accounting for an additional 12% of the variance. -In IDWS factor 1 (lack of energy) and IDWS factor 2 (danger) both predicted insomnia severity.	The internal consistency: Total scale $\alpha = 0.93$ lack of energy subscale: $\alpha = 0.94$ Danger subscale: $\alpha = 0.75$.	NR

(Continued)

Table 1 (Continued).

Author, year	Instruments	Sample	Validity	Reliability	Responsiveness and Interpretability
Tang & Harvey, 2004 ⁵⁶ Jansson-Frojmark et al 2011 ⁵⁵	Anxiety and Preoccupation about Sleep Questionnaire (APSQ)	110 university students (41 of them have insomnia as determined by PSQI score). Community dwelling sample from two counties in Sweden. -Classified according to sleep patterns into: 1. Insomniacs 2. Poor sleepers 3. Normal sleepers	<i>Criterion/concurrent validity:</i> The correlation with PSQI ($r = 0.44$, $p < 0.0001$) The correlation with BAI ($r = 0.37$, $p < 0.0001$) This indicates that higher scorers on the APSQ were associated with higher scores on PSQI (poorer sleep quality) and BAI (worse anxiety). Frojmark et al, 2011 <i>Discriminant validity:</i> Frojmark et al, 2011 The 10 APSQ items, the total APSQ scale, and the two retained factors discriminated the three sleep status groups. In all the 10 items ($F = 97.6-245.12$, $p < 0.001$ in all instances), the total scale ($F = 296.99$, $p < 0.001$), and the two factors ($F = 215.60-328.29$, $p < 0.001$ in both instances), the insomnia disorder group had higher scores than the other two groups, and the poor sleepers scored higher than the normal sleepers. The between group effect sizes for the 10 items ranged from 0.18 and 0.35, for the total scale 0.39, and for the two factors 0.33 and 0.41. <i>Convergent validity:</i> The APSQ and its two factors were significantly related to the Pre-Sleep Arousal Scale -Cognitive (PSAS-C) at ($r = 0.45-0.52$, to the DBAS-10 at a moderate to good level ($r = 0.50-0.61$), to the HADS Scale -Anxiety at fair level ($r = 0.3-0.49$), and to the HADS -Depression at a fair level ($r = 0.34-0.40$). ASPQ with daytime parameters: APSQ and its two subscales were moderately associated with: ● Sleep-onset latency: $r = 0.28-0.34$ ● WASO: $r = 0.32-0.37$, TST: $r = 0.26-0.31$ ● Early morning awakening: $r = 0.27-0.30$ ● The APSQ and the two subscales were correlated with sleep quality ($r = 0.40-0.48$). ASPQ with daytime impairment: -APSQ and its two factors were significantly correlated with daytime impairment ($r = 0.41-0.56$) -Correlations when the impairment item is removed from the composite score: APSQ, $r = 0.53$; first subscale, $r = 0.53$; second subscale, $r = 0.38$.	<i>Internal consistency:</i> Tang & Harvey, 2004 Cronbach's alpha for total scale = 0.92 Frojmark et al, 2011 Cronbach's alpha coefficients were: For total APSQ scale = 0.93 For first factor = 0.91 For the second factor = 0.86	NR

Ellis et al. 2007 ⁵⁷	Sleep preoccupation scale (SPS)	Validity study: University students and community sample	<i>Criterion/concurrent validity:</i> - SPS with the Sleep Associated Monitoring Index (SAMI): - Good correlation between total scores of SPS and SAMI; Poor ($r = 0.67$, $p < 0.001$), average ($r = 0.58$, $p < 0.001$) and good ($r = 0.62$, $p < 0.001$) sleepers. <i>Construct/convergent validity:</i> SPS subscales with Global PSQI-score: $p < 0.001$ CBC subscale: $r = 0.37$ AC subscale: $r = 0.55$ <i>Discriminant validity:</i> [$F(2, 721) = 57.27$, $p < 0.001$], poor sleeper reported higher levels of preoccupation scores than average and good sleepers. The tool significantly differentiated between poor and good sleepers (as identified through the PSQI questionnaire), as poor sleepers reported more preoccupation than normal sleepers. [$t(454) = 8.78$, $p < 0.001$].	Internal consistency: $\alpha = 0.91$ (overall) For subscales: CBC=0.93 and AC=0.89	NR
Tan et al. 2016 ⁶⁰	Catastrophic thoughts about insomnia scale (CTIS)	University students	<i>Content validity:</i> Expert panel reviewed the scale. <i>Face validity</i> ($n=523$): The tool was given to participants to complete. <i>Criterion validity:</i> - The correlation between CTIS and PSQI scores was statistically significant, $r(137) = 0.643$, $p < 0.001$ - The correlation between scores on the CTIS and scores on the ISI was also statistically significant, $r(137) = 0.703$, $p < 0.001$ <i>Predictive validity:</i> - CTIS along with DBAS-16, IDWS, nBFI, and CESD-10³, age, gender, and length of education predicted only 51% of the variance in PSQI score $F(8, 128) = 17.07$, $p < 0.001$, $R^2 = 0.51$. CTIS, CESD-10³, and gender significantly predicted changes in the PSQI scores. - CTIS, DBAS-16, IDWS, nBFI, and CESD-10³ and age, gender, and level of education predicted around two-thirds of the changes seen in ISIS scores $F(8, 128) = 32.07$, $p < 0.001$, $R^2 = 0.67$. Each of the three instruments CTIS, the CESD-10³, and the IDWS independently predicted the changes in ISI scores. <i>Construct/convergent validity:</i> CTIS scores were significantly associated with DBAS-16 scores, $r(137) = 0.722$, $p < 0.001$, and IDWS scores, $r(137) = 0.753$, $p < 0.001$.	Internal consistent ($\alpha = 0.94$) Subscales: Helplessness: $\alpha = 0.84$ Rumination: $\alpha = 0.88$	NR
Blake & Gomez, 1998 ⁶³	Sleep Hygiene Self-test. (SHS)	Male war-zone veterans	<i>Criterion validity:</i> Sleep Hygiene self-test with Combat exposure scale and Mississippi scale ($r = 0.20$ and 0.10 respectively).	Internal consistency: Overall: $\alpha = 0.54$	Responsiveness: After 5 sessions of education on sleep disturbance management (1 hour/week), a significant change of 7.6 points scores was detected

(Continued)

Table 1 (Continued).

Author, year	Instruments	Sample	Validity	Reliability	Responsiveness and Interpretability
Martin et al, 2006 ⁶¹	Sleep Hygiene Index (SHI)	Psychology students	Criterion: The SHI was positively correlated ($p < 0.01$) with the inadequate sleep hygiene criteria identified by the American Sleep Disorders Association, 1990. (r values = 0.371 to 0.458). Construct: -A positive association was identified between SHI score and both ESS and PSQI scores r (599) = 0.244 and r (269) = 0.481 respectively with $p < 0.01$ -The SHI scores were significantly correlated with PSQI components scores ($p \leq 0.05$).	Internal consistency: $\alpha = 0.66$ Test-retest reliability (repeated over 4–5 weeks) r (139) = 0.71, $p < 0.01$	NR
Brown et al, 2002 ⁶²	Sleep Hygiene Awareness and Practice Scale (SHAPS)	University students ($n=124$)	Construct: Correlation with sleep quality (as rated by the PSQI): - Sleep hygiene awareness with PSQI rating: ($r = 0.21$) - Sleep hygiene practice with PSQI rating: ($r = 0.49$, $p < 0.012$) - Sleep hygiene practice with sleep hygiene awareness ($r = 0.30$, $p < 0.012$). - Variable sleep length, noise disturbance, going to bed thirsty, and worrying about the ability to fall asleep at bedtime were identified as significant predictors of sleep quality predictors ($R^2 = 0.24$, adjusted $R^2 = 0.22$, $F(1, 118) = 5.30$, $p = 0.023$).	Internal consistency: Cronbach's $\alpha = 0.78$ The caffeine knowledge and sleep-hygiene practice subscales: Cronbach's $\alpha = 0.55$ and $\alpha = 0.47$, respectively. Test retest (4 weeks later): The sleep-hygiene awareness activities and sleep-hygiene practice: ($r = 0.76$, $p < 0.001$ and $r = 0.74$, $p < 0.001$, respectively). The caffeine knowledge subsection had poor test-retest reliability ($r = 0.50$, $p < 0.001$).	NR

Grandner et al, 2017 ⁵	Sleep Practices and Attitudes Questionnaire (SPAQ)	General population (Age range: 18–80) (n=124)	<i>Face validity:</i> evaluated through group conversations (eg focus group) between individuals from the community and research participants. The participants provided their feedback on the content of the instrument and discussed some specific items. <i>Content validity:</i> questions derived from theoretical framework were discussed by a group of professionals in sleep medicine and community members to ensure the comprehensiveness and representativeness of items. <i>Concurrent/criterion validity:</i> - A significant correlation was identified between PSQI sleep duration and average sleep duration ($r = 0.53$; $p < 0.001$). - Sleepiness with the ESS ($r_s = 0.39$, $p < 0.001$). - Coping with acute insomnia correlation with the SHI was moderate ($r = 0.29$, $p < 0.001$). - Activities in bed subscale with the SHI ($r = 0.53$, $p < 0.001$). - Sleep environment with SHI ($r = -0.34$, $p < 0.001$). - Impact of external factors on sleep scores on this subscale significantly differed between good and poor sleepers ($p = 0.004$). - Sleep quality and global PSQI score: ($r = 0.36$, $p < 0.001$) <i>Construct validity:</i> SPAQ with DBAS: DBAS was significantly correlated with subscales 2 ($r = 0.18$, $p < 0.05$), 3 ($r = 0.28$, $p < 0.01$), 4 ($r = 0.31$, $p < 0.001$), 8 ($r = 0.23$, $p < 0.05$), 12 ($r = 0.30$, $p < 0.001$), 13 ($r = 0.45$, $p < 0.001$), 15 ($r = 0.30$, $p < 0.001$), and 16 ($r = 0.26$, $p < 0.01$).	Internal consistency: Cronbach's alpha range: 0.251–0.864 (for the different subscales).	NR
-----------------------------------	--	---	---	--	----

Abbreviations: MOS, Medical Outcomes Study; GHQ-12, General Health Questionnaire-12; CES-D Scale, Centre for Epidemiologic Studies Depression; FSS, Fatigue Severity Scale; SWLS, Satisfaction with Life Scale; DIMS, Disorders of Initiating and Maintaining Sleep; DOES, Disorders of Excessive Somnolence; SOL, Sleep Onset Latency; SF-12, Short Form 12; POMS, Profile of Mood States; NPV, Negative Predictive Value; PPV, Positive Predictive Value; BNSQ, Basic Nordic Sleep Questionnaire; ICC, Inter-Correlation Coefficient; SD, Standard Deviation; CBT-I, Cognitive Behavioral Therapy for insomnia; HADS, Hospital Anxiety and Depression Scale; GBSS+, Great British Sleep Survey extended; A-D schedule, this consists of the State-Trait Anxiety Inventory (STAI-X1) and the Depressive Questionnaire (DQ); OSAS, Obstructive Sleep Apnea Syndrome; INS, Insomnia; BSD, Behavioral Sleep Disorders; OSA, obstructive sleep apnea; CPAP, Continuous Positive Airway Pressure; MSLT, Multiple Sleep Latency Test; VAS, Visual Analogue Scale; EEG, Electroencephalogram; EOG, Electro-Oculogram; SIP, Sickness Impact Profile; SF-36, Short Form 36; TST, Total Sleep Time; WASO, Wake After Sleep Onset; WPAL-GH, Work Productivity and Activity Impairment-General Health Questionnaire; BAI, Beck Anxiety Inventory; CBC, cognitive/behavioral consequences; AC, Affective Consequences; nBFI, Neuroticism Subscale of the Big Five Inventory; CASQ, Cognitive-somatic Anxiety Questionnaire.

Table 2 Characteristics of Sleep Instruments Extracted from Selected Studies (n=112)

Instruments	Measurement Model			Burden (1)	Response Format (2)	Cultural and Language Adaptation	Population in Which the Tool Was Validated
	Number of Qs	Scale	Scoring				
The Bergen insomnia scale (BIS)	6	8-points scale representing the frequency (number of days/week) ranging from 0 to 7	Total= 0 to 42	NR	Self- report	English	NR
The Insomnia Screening Scale (ISS)	26	NR	Insomnia symptoms: 0–50 Daytime function: 0–35 Sleep environment: 0–25 Sleep opportunity 0–20	NR	Self- report	NR	-Community sample: (6–72 years) -Primary insomnia
Sleep Quality Questionnaire (SQQ)	10	5-point Likert-type scale (0–4) (0= strongly disagree, 4 = strongly agree).	Total: 0–40	20 minutes	Self-report	Japanese	Adults
Women's Health Initiative Insomnia Rating Scale (WHIRS)	5	5-points Likert scale (except 1 question 6 points Likert scale)	Total: 0–20	NR	Self-report	English	Postmenopausal women
Restorative sleep questionnaire (RSQ)	RSQ-daily: 11 items RSQ- weekly: 9 items	Likert scale 1–5, in which 1 corresponds with “not at all” and 5 with “completely”.	Total score: The average of the questionnaire items then converted to 0–100 scale. A minimum of 5 items must be completed.	NR	Self-report	English	NR

Pittsburg sleep quality index (PSQI)	19 items (+ 5 questions for bed partner (optional))	First 4 items (open ended), PSQI items use varying response categories that include recording usual bedtime, usual wake time, number of actual hours slept, and number of minutes to fall asleep, as well as forced-choice four-point Likert scale different options including ("not during the past month", "less than once a week", "once or twice a week" and "three or more times a week").	Global PSQI score = 0- 21 Question score = 0 to 3	Past month	5-10 minutes	Self-report	English French India Japanese Chinese Spanish Hungarian Persian Serbian Thai Czech Korean Kurdish Arabic Hebrew Brazilian Portuguese Portuguese Sinhala Italian Malay	-US College students -Pregnant women -Lower back pain -University students -Psychiatric disordered people - Cancer patients -Breast cancer -Fibromyalgia (the Spanish version validated) -Older adults with Generalized anxiety disorders -Community dwelling Ethiopian adults
Insomnia severity index (ISI)	7 items	5-point Likert scale	Total: 0 to 28	Last 2 weeks	<5 minutes	Self-report Clinician Significant others (usually spouse).	English German French Spanish Chinese Italian Turkish Persian Arabic Korean Danish	-Veterans with traumatic brain injury -Lower back pain -Cancer patients -Community-dwelling older people

(Continued)

Table 2 (Continued).

Instruments	Measurement Model			Burden (1)	Response Format (2)	Cultural and Language Adaptation	Population in Which the Tool Was Validated
	Number of Qs	Scale	Scoring				
The Jenkins Sleep Scale (JSS)	4 items	6-point Likert scale, reflecting the frequency of having symptoms in the past month: (Not at all = 0; 1–3 days = 1, 4–7 days = 2; 8–14 days = 3; 15–21 days = 4; 22–28 days = 5).	Total: 0–20 (Higher scores indicate worse sleep quality; score >4 indicates the presence of sleep problem).	NR	Self-report	English Turkish Portuguese	-Post cardiac surgery -Air traffic controllers - Patients with rheumatoid arthritis -Patients with Psoriatic arthritis -Ankylosing spondylitis -Shift workers -Patients with chest pain
Athens Insomnia scale (AIS)	AIS-8: 8 items AIS-5: 5 items	Each item is scored on a scale 0–3 with 0 indicating “no problem” and 3 indicating a “very serious problem”. An item is rated positive if the described problem's frequency was ≥ 3 times/week	Total score range: AIS-8= 0–24 AIS-5= 0–15	NR	Self-report Administered through an interviewer	English Greek Spanish Japanese (simplified version: 1 item)	-Psychiatric patients -Individuals with primary insomnia - Chronic pain patients
Minimal insomnia Symptoms Scale (MISS)	3 items	5-point Likert scale (0= no, 4= very severe problems)	Total= 0–12	NR	Self-report	English	-Elderly

Sleep Condition Indicator (SCI)	8 items	5-point scale (0–4)	Total: 0 to 32 -Scores can be transformed into a more intuitive 0–10 SCI range: by dividing the total by 3.2, or by using an online version with automated scoring.	Typical night in the last month	NR	Self-report	English French Swedish	Community dwelling adults (French population)
General Sleep Disturbance Scale (GSDS)	21 items	8-point Likert scale ranging from 0 (never) to 7 (every day); higher scores indicate more severely disturbed sleep.	Total: 0–147	Previous week	5 minutes or less	Self-report	English Chinese Korean Italian	-Female nurses -Parents of infants admitted to ICU
Daily Cognitive Communication and Sleep Profile (DCCASP)	7 items	7-point Likert scale (1=worst function and 7=best function)	Total: 7–49	Daily	NR	-Self-report -Therapist -Significant others	English French	NR
Arabic Scale of Insomnia (ASI)	12 items	5-point Likert scale (0= No, 4=very much)	Total 0–48	Last month	NR	Self-reported	English Arabic	NR
Maugeri Sleep Quality and Distress Inventory (MasQuDI-17)	17	3-points Likert scale (1=Never, 2= sometimes and 3= frequently)	Total: 17 to 51	NR	NR	Self-report	English	NR

(Continued)

Table 2 (Continued).

Instruments	Measurement Model				Burden (1)	Response Format (2)	Cultural and Language Adaptation	Population in Which the Tool Was Validated
	Number of Qs	Scale	Scoring	Recall Period				
Epworth sleepiness scale (ESS)	8	Measured on 4-point Likert scale (0–3, corresponding to never dose off and high chance of dosing respectively)	Total score (0–24).	Recent times	NR	-Self-report -Completed by roommate	Arabic English Ethiopian German Greek Chinese (Eastern China) Italian Turkish Spanish	-Patients with sleep disorders -Patients with neurological disorders -Patients with psychiatric disorders -Workers (truck drivers) -Obstructive sleep apnea patients -African American population -Truck drivers
Karolinska sleepiness scale (KSS)	1 item	10 points Likert scale 1 (extremely alert) to 9 (extremely sleepy, fighting sleep)	Total score: 7–9	5 minutes before answering the question	NR	NR	English Japanese Spanish (for Mexican population)	Registered nurses (2)

Functional outcomes of sleep questionnaire (FOSQ)	FOSQ-30: 30 items FOSQ-10: 10 items	FOSQ-30: Six-point Likert scale (how frequently an activity is performed, 0 = never did it, 5 = three or more times a week) Subscale score calculation: A response score of 0 = a N/A or missing response. Thus, the potential range of scores for any item is 1–4. Calculate the mean of the answered items with responses equal to or greater than 1 for each subscale. This is the weighted mean item total or subscale score.	Global score: FOSQ-30: The mean of the subscale scores and multiply that by the number of subscales for which there is a score. - Scores range: 5–20 FOSQ-10: Total: the mean of the subscale scores multiplied by 5. Subscale score: 1–3.	Performance on a typical day (or generally how the person feels during the day).	FOSQ-30: 15 minutes FOSQ-10: NR	Self-administered paper-and-pencil questionnaire at a fifth-grade reading level.	English FOSQ-30: Thai Norwegian Swedish FOSQ-10: Persian Mandarin Chinese Spanish	FOSQ-30: -Obstructive sleep apnea patients -Veterans with type 2 diabetes -Women with breast cancer (stages I–III) -Pregnant women FOSQ-10: Family members of critically ill infants
Occupational impact of sleep questionnaire (OISQ)	24 items	5-point Likert scale. Scores on each question range from 0 ("never/not applicable to 4 "all of the time").	Total: 0–96	Last month	NR	NR	English Dutch Persian	NR
The Sleep Functional Impact Scale (SFIS)	26 items	5-points Likert scale	Total: 26–130 ≥ 13 items must be completed	7 days	NR	Self-report	English (US) Spanish (US) French	NR
Sleep disturbance questionnaire (SDQ)	12 items	5-point Likert scale ("never true," "seldom true," "sometimes true," "often true," "very often true")	NR	Typical night with poor sleep	NR	NR	NR	NR
Hyperarousal Scale (H-scale)	26	Measured on a four-point Likert-type scale from 0 "not at all," 1 "a little," 2 "quite a bit" and 3 "Extremely".	The scale measures the summation score (HSUM); with 3 scores per item.	NR	NR	Self-report	English Swedish	-Insomniacs -Patients with hypersomnia

(Continued)

Table 2 (Continued).

Instruments	Measurement Model				Burden (1)	Response Format (2)	Cultural and Language Adaptation	Population in Which the Tool Was Validated
	Number of Qs	Scale	Scoring	Recall Period				
Dysfunctional Beliefs and Attitudes about Sleep Scale (DBAS)	-DBAS: 30 or 31 items -DBAS-16: 16 items -DBAS 10: 10 items	Place a mark on a 10 cm long line within poles marking "strongly agree" and "strongly disagree".	DBAS full and DBAS 10: Average score across the items	Not applicable	NR	Self-report	English Chinese Taiwanese German (DBAS-16) Persian	NR
Insomnia daytime worry scale (IDWS)	11 items	5-point Likert scale.	NR	Last week	NR	Self-report	English	NR
Anxiety and Preoccupation about Sleep Questionnaire (APSQ)	10 items	-Original (2004): 10 points Likert scale (1= Not true to 10=very true) -Modified (2011): 5-point Likert scale. (1 = strongly disagree, 5= strongly agree)].	Total: Original:10–100 Modified: 10 to 50	Original: past 3 days Modified: last month	NR	Self-report	English	NR
Sleep preoccupation scale (SPS)	22 items	7-point Likert scale (0= never to 6= all of the time)	Total: 0–132	NR	NR	Self-report	English	NR
Catastrophic thoughts about insomnia scale (CTIS)	18 items	7-points Likert scale Ranges from 0=strongly disagree to 6= strongly agree and 3= neutral	Total: 0–108	NR	NR	Self-report	English	NR
Sleep hygiene-self test (SHS)	30 items	Dichotomous response (yes or No)	Total score (Yes = 1 point and No=0 point). Every fifth item is scored in the reverse direction	Previous month	NR	Self-report	English	NR

Sleep hygiene index (SHI)	13-item	5-point scale ranging from 0 (never) to 4 (always).	Total score: 0–52 or 13–65 (higher score indicates worse sleep hygiene)	NR	NR	Self-report	English Persian Turkish Korean	-Individuals with chronic pain -Elderly -University students (Nigerian and Saudi)
Sleep Hygiene Awareness and Practice Scale (SHAPS)	50 items	It asks respondents to rate the behavior on a scale from behavior is very beneficial to sleep (1) to behavior is very harmful to sleep (7). A rating of 4= behavior has no effect - 1 point= correct, 2 points if omitted and 3 points if incorrect (original developers lack et al, 1987) -Subsection 2: Score range: 0–7 Respondents are asked to write y if the substance contains caffeine, n if it does not, and x if they have never heard of it. (brown et al, 2002).	Awareness: Subsection 1: Total: Scores can range from 13 to 39, Subsection 2: This score can range from 0 to 100. Practice: The total: scores can range from 0 to 133; higher scores indicate behaviors more indicative of bad sleep hygiene	NR	NR	Self-report	NR	Insomniacs Good sleepers University students
Sleep Practices and Attitudes Questionnaire (SPAQ)	151 items	The sleep duration subscale: number of hours of habitual sleep, and the sleep debt subscale: the percentage of difference between need and habitual sleep duration (with values < 0 reflecting more obtained than needed and values > 0 reflecting less obtained than needed). -All other subscales scores range from 0 to 1.	NR	NR	10 minutes	Self-report	English	NR

Note: NR: Not reported; (1): Administration time; (2): Individuals who need to complete the questionnaire.

Table 3 Quality Assessment of Extracted Sleep Instruments* (n=31)

Sleep Instrument	Reliability		Validity			Responsiveness ^b	Interpretability ^b
	Internal Consistency ^a	Test Re-test ^b	Content ^c	Construct ^b	Criterion ^b		
Instruments screening for insomnia symptoms							
BIS	+	+	N/A	+	N/A	N/A	N/A
ISS	+	N/A	N/A	N/A	±	N/A	+
SQQ	+	+	+	+	N/A	N/A	N/A
WHIIRS	+	+	+	-	N/A	N/A	N/A
RSQ	+	+	+	+	N/A	N/A	N/A
PSQI	+	+	N/A	+	+	N/A	+
ISI	+	N/A	+	+	+	+	+
JSS	+	-	N/A	-	N/A	N/A	N/A
AIS	+	+	N/A	N/A	N/A	N/A	+
MISS	+	+	N/A	N/A	+	+	+
SCI	+	N/A	+	+	+	N/A	+
GSDS	+	N/A	N/A	N/A	-	N/A	N/A
DCCASP	+	+	N/A	N/A	-	N/A	N/A
ASI	+	+	+	+	?	N/A	N/A
Instruments evaluating consequences of poor sleep							
MaSQul-17	+	N/A	N/A	±	N/A	N/A	N/A
ESS	+	+	N/A	N/A	+	+	N/A
KSS	N/A	N/A	N/A	N/A	+	N/A	N/A
FOSQ	+	+	+	+	+	+	N/A
OISQ	+	+	N/A	N/A	+	N/A	N/A
SFIS	+	N/A	+	+	N/A	N/A	N/A
SDQ	+	N/A	N/A	N/A	N/A	N/A	N/A
H-scale	N/A	N/A	N/A	N/A	-	N/A	N/A
Instruments assessing the cognitive aspect of insomnia							
DBAS	+	N/A	N/A	+	?	+	N/A
IDWS	+	N/A	N/A	N/A	+	N/A	N/A
APSQ	+	N/A	N/A	-	N/A	N/A	N/A
SPS	+	N/A	N/A	-	-	N/A	N/A
CTIS	+	N/A	+	+	+	N/A	N/A
Instruments measuring sleep hygiene							
SHS	-	N/A	N/A	N/A	-	+	N/A
SHI	-	+	N/A	-	-	N/A	N/A
SHAPS	+	+	N/A	-	N/A	N/A	N/A
SPAQ	-	N/A	+	-	-	N/A	N/A

Notes: *Based on a modified version of the quality criteria developed by Terwee et al.²³ ^aCronbach's alpha: + if 0.7–0.95; - if (< 0.7 or >0.95). ^bPearson's correlation coefficient (r): + if r≥0.70; - if r<0.70. ^cContent validity: + if reported. N/A: No information available to establish rating; ±, conflicting results in studies or inconsistent results across subscales? questionable/not rated.

Abbreviations: BIS, Bergen insomnia scale; ISS, The Insomnia Screening Scale; SQQ, Sleep Quality Questionnaire; WHIIRS, Women's Health Initiative Insomnia Rating Scale; RSQ, Restorative sleep questionnaire; PSQI, Pittsburgh sleep quality index; ISI, Insomnia severity Index; JSS, Jenkins Sleep Scale; AIS, Athens Insomnia Scale; MISS, Minimal Insomnia Symptom Scale; SCI, Sleep Condition Indicator; GSDS, General Sleep Disturbance Scale; DCCASP, Daily Cognitive-Communication and Sleep Profile; AIS, Arabic Scale of Insomnia; MaSQul-17, Mageri Sleep Quality and Distress Inventory; ESS, Epworth Sleepiness scale; KSS, Karolinska sleepiness scale; FOSQ, Functional outcomes of sleep questionnaire; OISQ, Occupational Impact of sleep questionnaire; SFIS, The Sleep Functional Impact Scale; SDQ, Sleep Disturbance Questionnaire; H-Scale, Hyperarousal scale; DBAS, Dysfunctional Beliefs and Attitudes about Sleep Scale; IDWS, Insomnia daytime worry scale; APSQ, Anxiety and Preoccupation about Sleep Questionnaire; SPS, Sleep preoccupation scale; CTIS, Catastrophic thoughts about insomnia scale; SHS, Sleep Hygiene Self-test; SHI, Sleep Hygiene Index; SHAPS, Sleep Hygiene Awareness and Practice Scale; SPAQ, Sleep Practices and Attitudes Questionnaire.

the Hyperarousal Scale (H-Scale)^{48,49} also assesses daytime alertness. Three instruments, the FOSQ,^{74,28} Occupational Impact of Sleep Questionnaire (OISQ),⁵⁰ and the Sleep Functional Impact Scale (SFIS),³¹ evaluate the effects of insomnia on overall functioning. Conversely, the Sleep Disturbance Questionnaire (SDQ)⁵¹ evaluates factors which contribute to poor sleep, and the Maugeri Sleep Quality and Distress Inventory (MaSQuDI-17)⁵² measures the emotional burden of insomnia.

Instruments Assessing Cognitive Aspects of Insomnia

These were in line with the ICD-10 criteria which identifies preoccupation with sleepiness and excessive worry about the consequences of insomnia during the day as one of the clinical features for insomnia.⁵³ Four of these instruments [Insomnia Daytime Worry Scale (IDWS),⁵⁴ Anxiety and Preoccupation about Sleep Questionnaire (APSQ),^{55,56} Sleep Preoccupation Scale (SPS)⁵⁷] evaluate the extent of worry about insomnia. The Dysfunctional Beliefs and Attitudes about Sleep Scale (DBAS) explores individuals' perceptions and beliefs regarding insomnia.^{51,58,59} While the Catastrophic Thoughts about Insomnia Scale (CTIS)⁶⁰ assess exaggerated thinking about insomnia and the severity of the consequences.

Instruments Assessing Sleep Hygiene

Out of the four instruments included under this category [Sleep Hygiene Index (SHI),⁶¹ Sleep Hygiene Awareness and Practice Scale (SHAPS),⁶² Sleep Hygiene Self-test (SHS)⁶³ and the SPAQ⁷⁵], only the SPAQ is a comprehensive instrument which evaluates sleep quality, quantity, consequences of poor sleep, and sleep hygiene. The SHS also evaluates the effects of interventions which are expected to improve sleep hygiene, and is the only sleep-hygiene instrument which has been validated for responsiveness. The SHAPS measures knowledge about sleep hygiene in addition to behaviors surrounding sleep.

Some of the sleep instruments which were validated in the 107 articles included in this review, such as the BIS,³⁹ Sleep Condition Indicator (SCI),⁶⁴ ISS, Athens Insomnia Scale (AIS)^{65,66} and Sleep Hygiene Index (SHI)⁶¹ were developed in accordance with international diagnostic criteria for insomnia such as the International Classification of Sleep Disorders 2nd edition (ICSD-II), International Classification of Diseases 10th edition (ICD-10), and the

Diagnostic and Statistical Manual of Mental Disorders (DSM). The majority of these instruments were validated among university students and in primary insomniacs.

Many of the identified sleep instruments have been translated and culturally adapted in variety of languages, such as the Pittsburgh Sleep Quality Index (PSQI),^{67–69} ISI, and ESS which are available in 18, 10, and 6 languages, respectively.

Quality Assessment

The majority of the instruments used in the 107 studies included in this review did not meet all of the eight criteria set by the SAC-MOT.²² However, the conceptual and measurement models were specified for all instruments. Additionally, the reliability of sleep measures was usually reported in the validation studies, particularly the internal consistency reliability was reported for all instruments, except for the H-scale and the KSS. The reported internal consistency values for the sleep and sleep hygiene instruments across the studies included in this systematic review ranged between 0.53 and 0.97.

As outlined in Table 3, the test-retest reliability was only reported for 15 of the 31 instruments. Validity measures were reported for almost all of the instruments identified (29 out of 31), of which criterion validity was the most commonly reported and content validity the least reported validity measure. The generalizability of the psychometric properties reported was examined for three instruments (ISS, ISI and AIS) in community samples and primary insomniacs. Only six instruments (MaSQuDI-17, ESS, FOSQ, DBAS, APSQ, SPS) provided evidence of discriminating capacity between healthy individuals and poor sleepers. As summarized in Table 1, some of the sleep instruments were tested for other validity measures such as incremental validity, diagnostic validity, and external validity.^{40,64–67,70,76} Additionally, the SHS was the only instrument that was assessed for responsiveness to change in sleep hygiene after interventions.⁴⁶ With the exception of SHAPS, all of these instruments had low internal consistency ($\alpha < 0.7$) and only the SHI and the SHAPS were evaluated for their reproducibility.^{75,61–63}

Discussion

This study identified 31 sleep instruments and described in detail their psychometric properties as described in 107 validation and cultural adaptation studies. In the validation studies included in this review, only the ISI, FOSQ and MISS were evaluated for all four psychometric properties

such as responsiveness, interpretability, reliability and validity.^{25,73-28,69,76} Because most validation studies try measuring the same construct in different settings, it was expected that the sleep instruments in these studies were mostly tested for their reliability. For the most part, the sleep instruments identified in this systematic review had good internal consistency with Cronbach's alpha values ≥ 0.7 , which is in line with the recommended threshold for adequate reliability reported in the literature.⁷⁷ However, overall lower Cronbach's alpha values were reported for instruments measuring sleep hygiene. It has been suggested that the low internal consistency identified for sleep hygiene measures could be due to the definition of sleep hygiene, which consists of different factors (not necessarily related) that have the potential for negatively affecting sleep.⁶¹ Thus, items in sleep hygiene instruments may appear to be poorly related in the internal consistency test. This suggests that when assessing the psychometric properties of sleep hygiene instruments, the use of other validity and reliability measures is recommended to avoid depending on Cronbach's alpha values alone.

The validation studies included in this systematic review reported on the reproducibility of only 15 out of the 31 instruments identified. Mostly were assessed through the test-retest reliability method. However, the time frame between the two measurements was highly variable, ranging between 2 days and 6 months.^{27,69} There is still no consensus in the literature on the best time interval for reproducibility tests. Findings from studies that compared between different time intervals suggests that test results from an interval of 2 days and 2 months are similar to those resulting from a 2-weeks' interval.^{78,79}

Criterion validity was the most commonly reported validity measure in the studies included in this review. Concurrent validity (which compares the score of the instrument to that of a gold standard which assesses the same construct) was evaluated for 18 out of the 31 sleep instruments identified, whereas the predictive validity (which provides evidence of an instrument's ability to forecast a particular outcome in the future) was assessed only for 5 instruments. Considering that most of the validation studies included in this review were done on newly developed instruments, not prioritizing on assessing their predictive validity is understandable. Some of the validation studies such as those done on ISI, RSQ and PSQI, compared the scores generated by the instruments to non-subjective assessment tools such as the PSG.^{29,37,65} The

quality criteria for questionnaires developed by Terwee et al recommends providing the reasons for selecting a measure to be a gold standard in a validation study.²³ However, in the studies included in this systematic review, the choice of the instrument to be used as a comparator or as a gold standard was rarely justified. In situations where a gold standard measure is not available, construct validity could be used to assess the instrument's validity.⁸⁰ This is the case for some of the instruments included in this review, such as the RSQ, APSQ and SHAPS, which evaluate new sleep-related concepts for which gold standards are not available.

The diagnostic validity of an instrument is an important measure, particularly for clinicians, as it examines the extent to which the instrument can accurately differentiate between healthy individuals and insomniacs at a specific score.⁸¹ Diagnostic validity was reported for only five of the instruments identified in this review (specifically the PSQI, ISI, AIS, MISS and the SCI). Incremental validity (which describes the ability of an instrument to predict a variable of interest beyond what is possible by other existing instruments) was only tested for the Sleep Quality Questionnaire (SQQ).⁶² This measure is important as it facilitates the comparison between different instruments and provides evidence of which instrument is superior.

The selection of an insomnia instrument in practice is also affected by factors such as the length of the instrument, the time required for completion and the languages in which it is available. Around half of the instruments identified in this review consisted of ten questions or less, and those assessing the cognitive aspects of insomnia or sleep hygiene were usually longer. However, the duration required for completion was not commonly reported. Only six instruments were available in more than three languages, and only four in Arabic language.³⁴ The limited availability of instruments in Arabic language offers a valuable opportunity for researchers interested in sleep medicine in the Arab region to translate and validate these instruments.

Despite the comprehensiveness of this systematic review, the time elapsed between the initial database search and the publication of the results is an important limitation. To address this, a supplementary literature search using the same search strategy and databases was undertaken in April 2020, which identified additional validation studies and 7 new instruments including: the Insomnia Catastrophising Scale (ICS),^{82,83} the Single-item sleep quality scale (SQS),⁸⁴ the Lebanese insomnia

scale (LIS-18),⁸⁵ the Daytime Sleepiness Perception Scale-4 (DSPS-4),⁸⁶ the Indian Sleepiness Scale (InSS),⁸⁷ Athlete Sleep Behavior Questionnaire (ASBQ),⁸⁸ and the Non-restorative Sleep Scale (NRS)⁸⁹. These findings are indicative that the improvement in scientists' understanding of insomnia pathophysiology has been associated with a parallel increase in the development of instruments which measure different aspects of sleep dysfunction and related risk factors. Researchers also appear to be acknowledging the cultural effects on sleep behaviors and insomnia perception, as two of the newly developed instruments (LIS-18 and InSS) were developed to describe sleep from a non-Western perspective. Comparing between the psychometric properties of insomnia instruments across different populations is interesting and could be the focus of a future investigation. In addition, because this systematic review was focused on identifying instruments that assess the symptoms of insomnia, daytime consequences and those related to sleep hygiene, instruments assessing arousal before sleep only such as the Pre-Sleep Arousal Scale (PSAP)^{90,91} were excluded.

Conclusion

Different instruments are available for evaluating various aspects of sleep and sleep hygiene. The validity and reliability of most of these instruments have been tested and are well established. However, psychometric testing for several of the available sleep instruments remains incomplete, particularly responsiveness and interpretability. Our findings suggest that future studies should focus on reporting more psychometric measures to ensure the trustworthiness of the findings generated by these instruments. The number of sleep instruments available in languages other than English including Arabic are limited, indicating a need to translate and culturally adapt many of these instruments into various languages to be available for use in clinical practice and research in different populations.

Disclosure

The authors report no conflicts of interest in this work.

References

- Hirshkowitz M, Whiton K, Albert SM, et al. National sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Heal J Natl Sleep Found*. 2014;1(1):40–43. doi:10.1016/j.sleh.2014.12.010
- Roth T. Insomnia: definition, prevalence, etiology, and consequences. *J Clin Sleep Med*. 2007;3(5Suppl):S7–S10.
- Sateia MJ. International classification of sleep disorders-third edition: highlights and modifications. *Chest*. 2014;146(5):1387–1394. doi:10.1378/chest.14-0970
- Ferrie JE, Kumari M, Salo P, Singh-Manoux A, Kivimaki M. Sleep epidemiology-a rapidly growing field. *Int J Epidemiol*. 2011;40:1431–1437. doi:10.1093/ije/dyr203
- Wong WS, Fielding R. Prevalence of insomnia among Chinese adults in Hong Kong: a population-based study. *J Sleep Res*. 2011;20(1pt1):117–126. doi:10.1111/j.1365-2869.2010.00822.x
- Hiller RM, Johnston A, Dohnt H, Lovato N, Gradisar M. Assessing cognitive processes related to insomnia: a review and measurement guide for Harvey's cognitive model for the maintenance of insomnia. *Sleep Med Rev*. 2015;23:46–53. doi:10.1016/j.smrv.2014.11.006
- Pillai V, Drake CL. Sleep and repetitive thought: the role of rumination and worry in sleep disturbance. In: Babson KA, Feldne M, editors. *Sleep and Affect: Assessment, Theory, and Clinical Implications*. San Diego: Elsevier Academic Press; 2015:201–225. doi:10.1016/B978-0-12-417188-6.00010-4
- Jefferson CD, Drake CL, Scofield HM, et al. Sleep hygiene practices in a population-based sample of insomniacs. *Sleep*. 2005;28(5):611–615. doi:10.1093/sleep/28.5.611
- Gellis LA, Park A, Stotsky MT, Taylor DJ. Associations between sleep hygiene and insomnia severity in college students: cross-sectional and prospective analyses. *Behav Ther*. 2014;45(6):806–816. doi:10.1016/j.beth.2014.05.002
- Brick CA, Seely DL, Palermo TM. Association between sleep hygiene and sleep quality in medical students. *Behav Sleep Med*. 2010;8(2):113–121. doi:10.1080/15402001003622925
- Seun-Fadipe CT, Aloba OO, Oginni OA, Mosaku KS. Sleep hygiene index: psychometric characteristics and usefulness as a screening tool in a sample of Nigerian undergraduate students. *J Clin Sleep Med*. 2018;14(8):1285–1292. doi:10.5664/jcsm.7256
- Krystal AD. Sleep and psychiatric disorders: future directions. *Psychiatr Clin North Am*. 2006;29(4):1115–1130; abstract xi. doi:10.1016/j.psc.2006.09.001
- Baglioni C, Battagliese G, Feige B, et al. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. *J Affect Disord*. 2011;135(1):10–19. doi:10.1016/j.jad.2011.01.011
- Meng L, Zheng Y, Hui R. The relationship of sleep duration and insomnia to risk of hypertension incidence: a meta-analysis of prospective cohort studies. *Hypertens Res*. 2013;36(11):985–995. doi:10.1038/hr.2013.70
- Cappuccio FP, D'Elia L, Strazzullo P, Miller MA. Quantity and quality of sleep and incidence of type 2 diabetes: a systematic review and meta-analysis. *Diabetes Care*. 2010;33(2):414–420. doi:10.2337/dc09-1124
- Van de Water A, Holmes A, Hurley DA. Alternatives to polysomnography objective measurements of sleep for non-laboratory settings as alternatives to polysomnography – a systematic review. *J Sleep Res*. 2011;20(1pt2):183–200. doi:10.1111/j.1365-2869.2009.00814.x
- Martin JL, Hakim AD. Wrist actigraphy. *Chest*. 2011;139(6):1514–1527. doi:10.1378/chest.10-1872
- Buyse DJ, Yu L, Moul DE, et al. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. *Sleep*. 2010;33(6):29–33. doi:10.1093/sleep/33.6.781
- Klingman KJ, Jungquist CR, Perlis ML. Questionnaires that screen for multiple sleep disorders. *Sleep Med Rev*. 2016;1–8. doi:10.1016/j.smrv.2016.02.004
- Ji X, Liu J. Subjective sleep measures for adolescents: a systematic review. *Child Care Heal Dev*. 2016;42(6):825–839. doi:10.1111/cch.12376
- Devine EB, Hakim Z, Green J. A systematic review of patient-reported outcome instruments measuring sleep dysfunction in adults. *Pharmacoeconomics*. 2005;23(9):889–912. doi:10.2165/00019053-200523090-00003
- Lohr KN. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res*. 2002;11(3):193–205. doi:10.1023/A:1015291021312

23. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34–42. doi:10.1016/j.jclinepi.2006.03.012
24. Akerstedt T. Subjective and objective sleepiness in the active individual. *Intern J Neurosci.* 1990;52:29–37. doi:10.3109/00207459008994241
25. Broman JE, Smedje H, Mallon L, Hetta J. The minimal insomnia symptom scale (MISS): a brief measure of sleeping difficulties. *UPS J Med Sci.* 2008;113(2):131–142. doi:10.3109/2000-1967-221
26. Westergren A, Broman JE, Hellstrom A, Fagerstrom C, Willman A, Hagell P. Measurement properties of the minimal insomnia symptom scale as an insomnia screening tool for adults and the elderly. *Sleep Med.* 2015;16(3):379–384. doi:10.1016/j.sleep.2014.10.016
27. Jenkins CD, Stanton BA, Niemczyk SJ, Rose RM. A scale for the estimation of sleep problems in clinical research. *J Clin Epidemiol.* 1988;41(4):313–321. doi:10.1016/0895-4356(88)90138-2
28. Weaver TE, Chasens ER, Ratcliffe SJ. Development of the FOSQ-10: a short version of the functional outcomes of sleep questionnaire. *Sleep.* 2009;32(7):915–919. doi:10.1093/sleep/32.7.915
29. Bastien CH, Vallières A, Morin CM. Validation of the insomnia severity index as an outcome measure for insomnia research. *Sleep Med.* 2001;2(4):297–307. doi:10.1016/S1389-9457(00)00065-4
30. Morin CM, Belleville G, Bélanger L, Ivers H. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep.* 2011;34(5):601–608.
31. Bell C, McLeod LD, Nelson LM, Fehnel SE, Zografos LJ, Bowers B. Development and psychometric evaluation of a new patient-reported outcome instrument measuring the functional impact of insomnia. *Qual Life Res.* 2011;20(9):1457–1468. doi:10.1007/s11136-011-9885-8
32. Suleiman KH, Yates BC, Berger AM, Pozehl B, Meza J. Translating the pittsburgh sleep quality index into Arabic. *West J Nurs Res.* 2010;32(2):250–268. doi:10.1177/0193945909348230
33. Suleiman KH, Yates BC. Translating the insomnia severity index into Arabic. *J Nurs Sch.* 2011;43(1):49–53. doi:10.1111/j.1547-5069.2010.01374.x
34. Ahmed AE. Validation of Arabic versions of three sleep surveys. *Qatar Med J.* 2014;2014(2):130–136. doi:10.5339/qmj.2014.20
35. Abdel-Khalek AM. The development and validation of the Arabic scale of insomnia (ASI). *Sleep Hypn.* 2008;10(1):3–10.
36. Levine DW, Kaplan RM, Kripke DF, Bowen DJ, Naughton MJ, Shumaker SA. Factor structure and measurement invariance of the women's health initiative insomnia rating scale (study a). *Psychol Assess.* 2003;15(2):123–136. doi:10.1037/1040-3590.15.2.123
37. Levine DW, Kripke DF, Kaplan RM, et al. Reliability and validity of the women's health initiative insomnia rating scale (study b). *Psychol Assess.* 2003;15(2):137–148. doi:10.1037/1040-3590.15.2.137
38. Drake CL, Hays RD, Morlock R, et al. Development and evaluation of a measure to assess restorative sleep. *J Clin Sleep Med.* 2014;10(7):733–741. doi:10.5664/jcsm.3860
39. Pallesen S, Bjorvatn B, Nordhus IH, Sivertsen B, Hjørnevik M, Morin CM. A new scale for measuring insomnia: the bergen insomnia scale. *Percept Mot Ski.* 2008;107(3):691–706. doi:10.2466/pms.107.3.691-706
40. Yeh ZT, Chiang RP, Kang SC, Chiang CH. Development of the insomnia screening scale based on ICSD-II. *Int J Psychiatry Clin Pr.* 2012;16(4):259–267. doi:10.3109/13651501.2011.640938
41. Fung CHL, Nguyen M, Moineddin R, Colantonio A, Wiseman-Hakes C. Reliability and validity of the daily cognitive-communication and sleep profile: a new instrument for monitoring sleep, wakefulness and daytime function. *Int J Methods Psychiatr Res.* 2014;23(2):217–228. doi:10.1002/mpr.1422
42. Lee KA. Self-reported sleep disturbances in employed women. *Sleep.* 1992;15(6):493–498. doi:10.1093/sleep/15.6.493
43. Lee SY. Validating the General Sleep Disturbance Scale among Chinese American parents with hospitalized infants. *J Transcult Nurs.* 2007;18(2):111–117. doi:10.1177/1043659606298502
44. Johns MW. A new method for measuring daytime sleepiness: the epworth sleepiness scale. *Sleep.* 1991;14(6):540–545. doi:10.1093/sleep/14.6.540
45. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep.* 1992;15(4):376–381. doi:10.1093/sleep/15.4.376
46. Johns MW. Sleepiness in different situations measured by the epworth sleepiness scale. *Sleep.* 1994;17(8):703–710. doi:10.1093/sleep/17.8.703
47. Reyner L, Horne J. Falling asleep whilst driving: are drivers aware of prior sleepiness? *Int J Leg Med.* 1998;111(3):120–123. doi:10.1007/s004140050131
48. Pavlova M, Berg O, Gleason R, Walker F, Roberts S, Regestein Q. Self-reported hyperarousal traits among insomnia patients. *J Psychosom Res.* 2001;51(2):435–441. doi:10.1016/S0022-3999(01)00189-1
49. Regestein QR, Dambrosia J, Hallett M, Murauski B, Paine M. Daytime alertness in patients with primary insomnia. *Am J Psychiatry.* 1993;150(October):1529–1534.
50. David B. *Sleep Quality and Daytime Functioning in Primary Insomnia: A Prospective Study.* Loughbrgh Univ Institutional Repos; 2008.
51. Espie CA, Inglis SJ, Harvey L, Tessler S. Insomniacs' attributions psychometric properties of the dysfunctional beliefs and attitudes about sleep scale and the sleep disturbance questionnaire. *J Psychosom Res.* 2000;48(2):141–148. doi:10.1016/S0022-3999(99)00090-2
52. Morrone E, Sguazzin C, Bertolotti G, et al. Development and validation of the Maugeri sleep quality and distress inventory (MaSQuDI-17). *PLoS One.* 2017;12(7):e0180743. doi:10.1371/journal.pone.0180743
53. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines.* Geneva: WHO; 1992.
54. Kallestad H, Hansen B, Langsrud K, Hjemdal O, Stiles TC. Psychometric properties and the predictive validity of the insomnia daytime worry scale: a pilot study. *Cogn Behav Ther.* 2010;39(2):150–157. doi:10.1080/16506070903115317
55. Jansson-Frojmark M, Harvey AG, Lundh LG, et al. Psychometric properties of an insomnia-specific measure of worry: the anxiety and preoccupation about sleep questionnaire. *Cogn Behav Ther.* 2011;40(1):65–76. doi:10.1080/16506073.2010.538432
56. Tang NKY, Harvey AG. Correcting distorted perception of sleep in insomnia: a novel behavioural experiment? *Behav Res.* 2004;42(1):27–39. doi:10.1016/S0005-7967(03)00068-8
57. Ellis J, Mitchell K, Hogg H. Sleep preoccupation in poor sleepers: psychometric properties of the sleep preoccupation scale. *J Psychosom Res.* 2007;63(6):579–585. doi:10.1016/j.jpsychores.2007.07.011
58. Edinger JD, Wohlgemuth WK. Psychometric comparisons of the standard and abbreviated DBAS-10 versions of the dysfunctional beliefs and attitudes about sleep questionnaire. *Sleep Med.* 2001;2(6):493–500. doi:10.1016/S1389-9457(01)00078-8
59. Chung KF, Ho FYY, Yeung WF. Psychometric comparison of the full and abbreviated versions of the dysfunctional beliefs and attitudes about sleep scale. *J Clin Sleep Med.* 2016;12(6):821–828. doi:10.5664/jcsm.5878
60. Tan LO, Hadjistavropoulos T, MacNab YC. The catastrophic thoughts about insomnia scale (CTIS): development and validation. *Cognit Ther Res.* 2016;41(1):143–154. doi:10.1007/s10608-016-9804-y
61. Mastin DF, Bryson J, Corwyn R. Assessment of sleep hygiene using the sleep hygiene index. *J Behav Med.* 2006;29(3):223–227. doi:10.1007/s10865-006-9047-6
62. Brown FC, Buboltz W Jr., Soper B. Relationship of sleep hygiene awareness, sleep hygiene practices, and sleep quality in university students. *Behav Med.* 2002;28(1):33–38. doi:10.1080/08964280209596396

63. Blake DD, Gómez MH. A scale for assessing sleep hygiene: preliminary data. *Psychol Rep.* 1998;83(3 Pt 2):1175–1178. doi:10.2466/PRO.83.7.1175-1178
64. Espie CA, Kyle SD, Hames P, Gardani M, Fleming L, Cape J. The sleep condition indicator: a clinical screening tool to evaluate insomnia disorder. *BMJ Open.* 2014;4(3). doi:10.1136/bmjopen-2013-004183
65. Soldatos CR, Dikeos DG, Paparrigopoulos TJ. Athens insomnia scale: validation of an instrument based on ICD-10 criteria. *J Psychosom Res.* 2000;48(6):555–560. doi:10.1016/S0022-3999(00)00095-7
66. Soldatos CR, Dikeos DG, Paparrigopoulos TJ. The diagnostic validity of the Athens Insomnia Scale. *J Psychosom Res.* 2001;55(3):263–267. doi:10.1016/S0022-3999(02)00604-9
67. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 1989;28(2):193–213. doi:10.1016/0165-1781(89)90047-4
68. Grandner MA, Kripke DF, Yoon IY, Youngstedt SD. Criterion validity of the pittsburgh sleep quality index: investigation in a non-clinical sample. *Sleep Biol Rhythms.* 2006;4(2):129–136. doi:10.1111/j.1479-8425.2006.00207.x
69. 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.* 2002;53(3):737–740. doi:10.1016/S0022-3999(02)00330-6
70. Kato T. Development of the sleep quality questionnaire in healthy adults. *J Heal Psychol.* 2013;19(8):977–986. doi:10.1177/1359105313482168
71. Ohayon MM, Smirne S. Prevalence and consequences of insomnia disorders in the general population of Italy. *Sleep Med.* 2002;3(2):115–120. doi:10.1016/S1389-9457(01)00158-7
72. Jerlock M, Gaston-Johansson F, Kjellgren KI, Welin C. Coping strategies, stress, physical activity and sleep in patients with unexplained chest pain. *BMC Nurs.* 2006;5(1):7. doi:10.1186/1472-6955-5-7
73. Naseri-moadeli A, Sekine M, Kumari M, Chandola T, Marmot M, Kagamimori S. Association of sleep quality and free time leisure activities in japanese and british civil servants. *J Occup Health.* 2005;47(5):384–390. doi:10.1539/joh.47.384
74. Weaver E, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep.* 1997;20(10):835–843.
75. Grandner MA, Jackson N, Gooneratne NS, Patel NP. The development of a questionnaire to assess sleep-related practices, beliefs, and attitudes. *Behav Sleep Med.* 2014;12(2):123–142. doi:10.1080/15402002.2013.764530
76. Morin CM, Belleville G, Belanger L, Ivers H. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep.* 2011;34(5):601–608.
77. Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD, Sloan JA. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health.* 2007;10(SUPPL.2):94–105. doi:10.1111/j.1524-4733.2007.00272.x
78. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. *J Clin Epidemiol.* 2003;56(8):730–735. doi:10.1016/S0895-4356(03)00084-2
79. Chmielewski M, Watson D. What is being assessed and why it matters: the impact of transient error on trait research. *J Pers Soc Psychol.* 2009;97(1):186–202. doi:10.1037/a0015618
80. Bannigan K, Watson R. Reliability and validity in a nutshell. *J Clin Nurs.* 2009;18(23):3237–3243. doi:10.1111/j.1365-2702.2009.02939.x
81. Smith GE, Cerhan JH, Ivnik RJ. Diagnostic validity. In: *Clinical Interpretation of the WAIS-III and WMS-III*. Minnesota: Elsevier Science (USA). 2003;273–300.
82. Jansson-Fröjmark M, Harvey AG, Flink IK. Psychometric properties of the Insomnia Catastrophizing Scale (ICS) in a large community sample. *J of Sleep Res.* 2012;21(Suppl1):168. doi:10.1080/16506073.2019.1588362
83. Jansson-Fröjmark M, Harvey AG, Flink I. Psychometric properties of the Insomnia Catastrophizing Scale (ICS) in a large community sample. *Cogn Behav Pract.* 2020;49(2):120–136. doi:10.1080/16506073.2019.1588362
84. Snyder E, Cai B, DeMuro C, Morrison MF, Ball W. A new single-item sleep quality scale: results of psychometric evaluation in patients with chronic primary insomnia and depression. *J Clin Sleep Med.* 2018;14(11):1849–1857. doi:10.5664/jcsm.7478
85. Hallit S, Sacre H, Haddad C, et al. Development of the Lebanese insomnia scale (LIS-18): a new scale to assess insomnia in adult patients. *BMC Psychiatry.* 2019;19(1):1–11. doi:10.1186/s12888-019-2406-y
86. Marques DR, Gomes AA, de Azevedo MHP. DSPS-4: a brief measure of perceived daytime sleepiness. *Curr Psychol.* 2019;38(2):579–588. doi:10.1007/s12144-017-9638-0
87. Venkatnarayan K, Krishnaswamy UM, Ramachandran P, Devaraj U. Validation of the Indian sleepiness scale to assess excessive daytime somnolence. *Sleep Med.* 2019;64:S404. doi:10.1016/J.SLEEP.2019.11.1124
88. Driller MW, Mah CD, Halson SL. Development of the athlete sleep behavior questionnaire: a tool for identifying maladaptive sleep practices in elite athletes. *Sleep Sci.* 2018;11(1):37–44. doi:10.5935/1984-0063.20180009
89. Wilkinson K, Shapiro C. Development and validation of the Nonrestorative Sleep Scale (NRSS). *J Clin Sleep Med.* 2013;9(9):929–937. doi:10.5664/jcsm.2996
90. Nicassio FM, Mendlowitz DR, Fussell JJ, Petras L. The phenomenology of the pre-sleep state: the development of the pre-sleep arousal scale *. *Behav Res Ther.* 1985;23(3):263–271. doi:10.1016/0005-7967(85)90004-X
91. Jansson-Fröjmark M, Norell-Clarke A, Jansson-Fröjmark M, Norell-Clarke A. Psychometric properties of the Pre-Sleep Arousal Scale in a large community sample. *J Psychosom Res.* 2012;72(2):103–110. doi:10.1016/j.jpsychores.2011.10.005

Nature and Science of Sleep

Publish your work in this journal

Nature and Science of Sleep is an international, peer-reviewed, open access journal covering all aspects of sleep science and sleep medicine, including the neurophysiology and functions of sleep, the genetics of sleep, sleep and society, biological rhythms, dreaming, sleep disorders and therapy, and strategies to optimize healthy sleep.

Submit your manuscript here: <https://www.dovepress.com/nature-and-science-of-sleep-journal>

Dovepress

The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.