Assessing Quality of Sleep in Patients with Rheumatoid Arthritis

GEORGE A. WELLS, TRACY LI, JOHN R. KIRWAN, JOAN PETERSON, DANIEL ALETAHA, MAARTEN BOERS, BARRY BRESNIHAN, MAXIME DOUGADOS, LEANNE IDZERDA, JO NICKLIN, MARIA SUAREZ-ALMAZOR, VIVIAN WELCH, and PETER S. TUGWELL

ABSTRACT. We sought to identify instruments assessing sleep quality that measure the domains of sleep applicable to rheumatoid arthritis (RA) patients and are feasible to use and have appropriate reliability, validity, and responsiveness properties. A systematic review of sleep instruments was conducted. In particular, domains related to sleep that were assessed in the instruments were identified and evaluated. Feasibility characteristics and psychometric properties of instruments were reviewed. At OMERACT 9, the preparatory work was described at the plenary session of the Patient Perspective Workshop, and the tasks of 3 breakout groups in ranking and scoring the domains and sleep instruments were outlined. Each breakout group considered different aspects of sleep: sleep domains, feasibility, and psychometric properties. The rapporteur for each breakout group reported back to the plenary on the domains and sleep instruments that achieved the highest rank/score. The systematic review identified 45 sleep instruments of interest. Based on these instruments, 14 domains of sleep were identified. The top ranked domains were: Sleep Adequacy (1), Sleep Maintenance (2), Sleep Initiation (3) and Daytime Functioning (4). The top ranked instruments on feasibility were: Athens Insomnia Scale (2.3), Medical Outcome Study (MOS) Sleep (4.0), Insomnia Severity Index (4.9), and Women's Health Insomnia Rating Scale (5.5). The highest scored instruments on psychometric properties were: Athens Insomnia Scale (13.6), Sleep Assessment Questionnaire (13), Pittsburgh Sleep Diary (12), and MOS Sleep (11). Sleep domains have been reviewed, and several sleep instruments have been identified. These instruments should be considered for use in planned clinical trials of RA patients to assess their applicability. (J Rheumatol 2009;36:2077–86; doi:10.3899/jrheum.090362)

Key Indexing Terms:
RHEUMATOID ARTHRITIS SLEEP QUALITY DOMAINS PSYCHOMETRIC PROPERTIES

From the Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada; Bristol-Myers Squibb, Princeton, New Jersey, USA; University of Bristol, Bristol, UK; Ottawa Health Research Institute, Ottawa Hospital, Ottawa, Canada; Medical University of Vienna, Vienna, Austria; Department of Clinical Epidemiology and Biostatistics, VU University Medical Centre, Amsterdam, Netherlands; Conway Institute of Biomolecular and Biomedical Research, University College Dublin and St Vincent's University Hospital, Dublin, Ireland; Paris-Descartes University, Paris, France; Institute of Population Health, University of Ottawa, Ottawa, Canada; Department of Medicine, Baylor College of Medicine, Houston, Texas, USA.

G.A. Wells, PhD, Department of Epidemiology and Community Medicine, University of Ottawa; T. Li, PhD, Bristol-Myers Squibb; J.R. Kirwan, MD, University of Bristol; J. Peterson, BSc, Ottawa Health Research Institute, Ottawa Hospital; D. Aletaha, MD, Medical University of Vienna; M. Boers, MD, Msc, PhD, Department of Clinical Epidemiology and Biostatistics, VU University Medical Centre; B. Bresnihan, MD, Conway Institute of Biomolecular and Biomedical Research, University College Dublin and St Vincent's University Hospital; M. Dougados, MD, Paris-Descartes University; L. Idzerda, BSc, Institute of Population Health, University of Ottawa; J. Nicklin, RN, MSc, University of Bristol; M. Suarez-Almazor, MD, MSc, Department of Medicine, Baylor College of Medicine; V. Welch, MSc, Institute of Population Health, University of Ottawa; P.S. Tugwell, MD, MSc, Institute of Population Health, University of Ottawa.

Supported in part by an unrestricted research grant-in-aid from Bristol-Myers Squibb.

Address correspondence to Dr. G.A. Wells, Department of Epidemiology and Community Medicine, University of Ottawa, 451 Smyth Rd., Ottawa, Ontario, KIH 8M5, Canada. E-mail: gawells@ottawaheart.ca

Patient reported outcomes provide an assessment of a patient's health, well-being, and treatment from the patient's perspective. Sleep quality and fatigue have been identified at different OMERACT meetings as important aspects of the health and well-being of patients with arthritis. In particular, at the OMERACT 6 workshop for developing an operational definition of low disease activity state for rheumatoid arthritis (RA), the patient group emphasized fatigue and sleep as important issues in RA[1]; and at a patient perspective workshop at OMERACT 7 the question of assessing outcomes of treatment for arthritis from the perspective of those who experience the disease themselves was addressed with a particular emphasis placed on fatigue[2]. The focus here is on sleep in patients with RA. Individuals with a variety of common medical illnesses including arthritis frequently experience sleep disturbances. It is recognized that medical illnesses can adversely affect sleep quality, and that pain, infection, and inflammation can induce symptoms of excessive daytime sleepiness and fatigue[3-6]. In particular, this is true for patients with RA[7-10].

Questionnaires are often the instrument of choice to assess sleep, and in using a particular instrument attention must be given to 3 aspects: the domains of sleep that are

evaluated, the feasibility of completing the questionnaire, and the psychometric or measurement properties of the instrument.

First, various domains of sleep have been identified and classification systems for sleep disorders derived. For example, the Diagnostic Classification of Sleep and Arousal Disorders[11] grouped sleep disorders into 4 major categories based on the primary symptom: insomnias (initiating and maintaining sleep), excessive sleepiness, sleep-wake schedule and parasomnias (dysfunctions of sleep, sleep stages, or partial arousals). The International Classification of Sleep Disorders[12] included insomnias, sleep-related breathing disorders, hypersomnia of central origin not due to circadian rhythm, sleep-related breathing or other causes of disturbed nocturnal sleep disorders, circadian rhythm sleep disorders, parasomnias, and sleep-related movement disorders. Hays and Stewart in The Medical Outcomes Study[13] identified domains: initiation, maintenance, quantity, perceived adequacy, somnolence, respiratory impairments, regularity, sleep stage disorders and use of sleep medications.

Second, feasibility relates to the efficiency of the administration of the instrument and the resulting burden of completing the instrument. In particular, this includes both the number of questions and the difficulty in answering the questionnaires based on the questionnaire format, response key, and language level. It is generally known that response rates and validity of the answers are directly related to feasibility[14]. Ideally, the instrument should have a short administration time, low reading level required, and be easily understood.

Third, the psychometric properties of an instrument of interest refer to the reliability, validity, and sensitivity of the instrument. Reliability is concerned with whether the instrument consistently measures the characteristic of interest, validity relates to whether the instrument measures what it is supposed to measure, and sensitivity to change is concerned with whether the instrument can detect small but clinically important changes. These properties are of particular importance when subjective reports of health status is one of the primary outcomes of the trial.

For properly assessing sleep for patients with RA, 3 key aspects of any sleep instrument need to be considered: the domains, feasibility, and psychometric properties. In terms of the OMERACT filter: truth relates to the domains assessed (content validity) and psychometric properties of validity and reliability; feasibility is directly related to administrative burden and applicability; and discrimination relates to the psychometric property of sensitivity or responsiveness. The first step is a systematic review of the literature for potential sleep instruments that could be used and then attaining consensus on which instruments should be further considered. The objective of our workshop was to identify instruments assessing sleep quality that measure domains of sleep applicable to RA patients and are feasible to use and have appropriate reliability, validity and responsiveness properties.

PREPARING FOR THE PATIENT PERSPECTIVE WORKSHOP ON SLEEP

In preparing for OMERACT 9, the working group met periodically by teleconference and E-mail in addition to in-person meetings at the American College of Rheumatology and European League Against Rheumatism conferences in 2007. A systematic literature review of instruments designed to assess various aspects of sleep was conducted in January 2007, and during 2007 these instruments were evaluated on their response characteristics, psychometric properties, and domains of sleep assessed. The deliverables for OMERACT 9 were to present the results of the systematic literature review on sleep instruments and their truth and feasibility of use in RA. The objective for OMERACT 9 was to select candidate instruments based on truth, discrimination, and feasibility that measure sleep domains of interest.

Systematic review of sleep instruments. In conducting the systematic review the methodology of the Cochrane Collaboration was adhered to and the following steps were undertaken: a comprehensive literature search was conducted (keywords and MeSH terms: sleep, insomnia, sleep disorders, questionnaires, interviews, health surveys, psychometrics, health status, quality of life); citations and articles were selected using predefined criteria by 2 independent reviewers; information on the instruments was extracted from the articles using 2 independent reviewers; characteristics of the instruments were summarized including format properties, number of items, response format, timeline, and psychometric properties (reliability, validity, responsiveness). The literature search included: Medline (1966 to January 2007), PsychINFO (1806 to January 2007), Web-based databases (MAPI Research Institute and Educational Testing Service Test Collection), sleep assessment textbook chapters, bibliographies of sleep research, and review articles. Self-report instruments designed to assess sleep and sleep disorders in adults were selected. Instruments developed to measure sleep disruption secondary to other medical conditions (e.g., Parkinson's disease, sleep apnea) were excluded, with the exception of chronic pain.

The search resulted in 3751 citations from Medline (1966 to January 2007) and 174 citations from PsychINFO (1806 to January 2007). After applying the selection criteria, 45 instruments were identified that assessed a variety of domains related to sleep (Table 1)[15-64]. In particular, the domains related to sleep that were assessed in the sleep instruments were identified and summarized, their applicability to chronic diseases, and in particular RA, were evaluated, and the psychometric properties and feasibility aspects of the instruments were reviewed.

The various domains related to sleep that were assessed

Table 1. Sleep instruments ascertained in the systematic review.

Abbreviation	Title and Reference
ATSI	Accumulated Time with Sleepiness Scale[15]
AIS	Athens Insomnia Scale[16]
BNSQ	Basic Nordic Sleep Questionnaire[17]
BSIQ	Brock Sleep and Insomnia Questionnaire[18]
DSD	Daily Sleep Diary[19]
Dutch SDQ	Dutch Sleep Disorders Questionnaire[20]
DBAS	Dysfunctional Beliefs and Attitudes About Sleep Questionnaire[21]
ESS	Epworth Sleepiness Scale[22,23]
Espie SDQ	Espie Sleep Disturbance Questionnaire[24]
FOSQ	Functional Outcomes of Sleep Questionnaire[25]
GCTI	Glasgow Content of Thoughts Inventory[26]
HS	Hyperarousal Scale[27]
ISI	Insomnia Severity Index[28]
Jenkins SEQ	Jenkins Sleep Evaluation Questionnaire[29]
KSD	Karolinska Sleep Diary[30]
KSS	Karolinska Sleepiness Scale[31]
Leeds SEQ	Leeds Sleep Evaluation Questionnaire[32]
MOS	MOS Sleep[33]
PSS	Pictorial Sleepiness Scale[34]
PghSD	Pittsburgh Sleep Diary[35]
PSQI	Pittsburgh Sleep Quality Index[36]
PSI	Post Sleep Inventory[37]
PSAS	Pre-sleep Arousal Scale[38]
QOLI	Quality of life & Insomnia[39,40]
RSS	Resistance to Sleepiness Scale[41]
RDSS	Rotterdam Daytime Sleepiness Scale[42]
SLEEP-50	Sleep-50 Questionnaire[43]
SAQ	Sleep Assessment Questionnaire[44,45]
SBSR	Sleep Behaviour Self Rating Scale[46]
SBS	Sleep Beliefs Scale[47]
SDsQ	Sleep Disorders Questionnaire[48]
SDsQ	Sleep Dissatisfaction Questionnaire[49]
SEI	Sleep Effects Index[50]
SEQ	Sleep Evaluation Questionnaire[51]
SHI	Sleep Hygiene Index[52]
SAMI	Sleep Associated Monitoring Index[53]
SQS	Sleep Quality Scale[54]
SQ	Sleep Questionnaire[55]
SSES	Sleep Self Efficacy Scale[56]
STQ	Sleep Timing Questionnaire[57]
SWAI	Sleep Wake Activity Inventory[58]
SSS	Stanford Sleepiness Scale[59]
SMHSQ	St. Mary's Hospital Sleep Questionnaire[60,61]
VSH Sleep Scale	Verran and Snyder-Halpern Sleep Scale[62]
WHIIRS	Women's Health Initiative Insomnia Rating Scale[63,64]

in the sleep instruments identified in the systematic review were itemized and summarized (Table 2). Fourteen domains were identified and presented at the EULAR 2007 conference. At a meeting of the working group at EULAR 2007, the applicability of these domains to chronic diseases, and in particular RA, was evaluated and confirmed. Also, the response characteristics and psychometric properties of the instruments were identified and summarized and, in preparation for OMERACT 9, a Delphi process reduced the number of instruments for consideration at OMERACT to 15 instruments. Selection of the instruments followed a similar process that would be used at the OMERACT meeting. The response characteristics of the instruments are summarized in Table 3. The number of items typically ranged from 1 to 30 items, the response format was usually a Likert scale (4 or 5 point) or visual analog scale, and the timeline ranged from "recent" to 3 months. Most of the instruments were multidomain, and a summary of their psychometric properties based on the primary report of the instrument is provided in Appendix A.

BREAKOUT GROUP SESSIONS FOR THE PATIENT PERSPECTIVE WORKSHOP ON SLEEP

At OMERACT 9, the session on sleep was part of the Patient Perspective Workshop that was designed to consider: a Patient Core Set, Sleep, Effective Consumer, and Psychological and Educational Interventions. At the plenary session for the Patient Perspective Workshop, the preparatory work was described and the tasks of 3 breakout groups for sleep were outlined. Each of the breakout groups considered different aspects of sleep: sleep domains, feasibility, and psychometric properties. The rapporteur for each breakout group reported back to the Patient Perspective Workshop on the deliberations of their group. They described the process and any key points raised during the breakout session and provided a summary of the rankings and scorings.

Sleep domains. For this breakout group, a deck of 14 cards was given to each participant. On each card was the identification of a domain related to sleep and a brief description (Table 2), and the participant was to reorder the cards from the most important to the least important domain based on their opinion. Although the domain descriptions were self-explanatory and were in lay language, if needed, the Chair of the breakout group could briefly review the domains. Once the task was completed, each participant returned the card deck ordered from the most important to the least important domain. In reporting back to the Patient Perspective Workshop the following were the 4 highest ranked domains: 1. Sleep Adequacy; 2. Sleep Maintenance; 3. Sleep Initiation; 4. Daytime Functioning.

Feasibility. A package of 15 sheets was given to each participant. On each sheet the identification of the instrument and a summary of the format of the instrument were provided. In addition there was a description of the instrument taken from the primary publication, which could vary from the instrument itself to a listing of the items in the instrument to a simple text description. If needed the Chair of the breakout group could review the "feasibility" component of OMERACT filter of "Truth, Discrimination and Feasibility." The participant was to reorder the sheets from the most feasible to the least feasible to use based on their opinion. In reporting back to the Patient Perspective Workshop the following were the 4 highest ranked sleep instruments based on

Table 2. Sleep domains derived from the sleep instruments in the systematic review.

Sleep Domain	Lay Description
Sleep initiation	The ability to fall asleep. The time required to fall asleep
Sleep maintenance	The ability to stay asleep all through the night or to get back to sleep if awakened
Sleep adequacy	Getting sufficient quality and quantity of sleep so as to feel rested on awakening.
Daytime sleepiness	Feeling sleepy during the day or having difficulty staying awake during quiet daytime activities
Sleep quantity	Hours of nighttime sleep
Sleep regularity	The extent to which sleep onset and arising are consistent from day to day
Sleep related behaviors	Behaviors carried out both during the day and before bed that would affect the ability to sleep. For example: daytime napping, shift work, meals, exercise, travel time zones, caffeine, alcohol, tobacco, and presleep activities that are either quiet or stimulating
Sleep related beliefs	Beliefs about one's own ability to sleep and beliefs about sleep in general
Physical comfort	Sleeping conditions such as room temperature, noise, light, bed partner, mattress, pillow or sleeping position. Physical conditions such as pain or muscle cramps which would interfere with comfort
Breathing problems	Problems at night with snoring, snorting, gasping, breath cessation, or shortness of breath
Sleep stage disorders	Sleepwalking, nightmares, bedwetting, teeth grinding
Anxiety/tension	Inability to unwind, relax, or turn off thoughts
Medication	Sleeping medications or medications taken for other conditions that would affect sleep
Daytime functioning	Ability to carry out work, leisure, household activities, and social relationships

Table 3. Response format of the selected sleep instruments from the systematic review.

Sleep Instrument	Characteristics
Athens Insomnia Scale	Timeline: The last month; 2 versions of scale available: AIS-8 (full scale version, consisting of 8 items relating both to sleep characteristics and daytime consequences) and AIS-5 relating to sleep characteristics only; 3 point rating scale
Daily Sleep Diary	Timeline: The previous night; 9 items; combination of short answer and 4 to 5 point ordinal scales
Dutch Sleep Disorders Questionnaire	Timeline: Not specified; 34 items; 4 point scale; developed from the 176 item Sleep Disorders Questionnaire
Dysfunctional Beliefs and Attitudes About Sleep Questionnaire	Timeline: Not specified; 30 items; VAS—strongly disagree to strongly agree
Epworth Sleepiness Scale	Timeline: "Recent times"; 8 items; Likert 4 point scale
Insomnia Severity Index	Timeline: Past 2 weeks; 7 items; six 5 point Likert and one multiple choice; available in 3 versions: self-administered, significant other, and clinician
Leeds Sleep Evaluation Questionnaire	Timeline: Comparison of sleep on medication to usual sleep; 10 items; 100 mm line analog scale
MOS Sleep	Timeline: Past 4 weeks; 12 items; 6 point Likert scale ranging from all of the time to most of the time
Pittsburgh Sleep Diary	Timeline: Bedtime portion pertains to current day; wake time portion pertains to previous night. Bedtime form 8 items; Wake time form 14 items; Variable format includes question and answer, circling the number of times a behaviour occurs, and three 10 cm VAS scales for subjective sleep quality, mood and alertness on awakening
Pittsburgh Sleep Quality Index	Timeline: Past month; 19 self-rated questions and 5 questions rated by bed partner (if available). Only self-rated questions are included in scoring; 4 point Likert (primarily) as well as some question and answer; 7 component scores (sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbance, use of sleeping medications, and daytime dysfunction) as well as a global sleep quality score
Sleep Assessment Questionnaire	Timeline: Not specified; 19 items; 5 point Likert scale
Sleep Disorders Questionnaire	Timeline: Not specified; 5 point Likert scale; 175 items
Sleep Dissatisfaction Questionnaire	Timeline: Not specified; 30 items; 5 point Likert
Stanford Sleepiness Scale	Timeline: The present; 7 statements from which subjects pick the one that best describes their state of sleepiness at that time
Women's Health Initiative Insomnia Rating Scale	Timeline: Past 4 weeks; 5 items; 5 point scale

feasibility: Athens Insomnia Scale, 2.3; MOS Sleep Measure, 4.0; Insomnia Severity Index, 4.9; Women's Health Insomnia Rating Scale, 5.5.

Psychometric properties. A package of 15 sheets was given to each participant. On each sheet the identification of the instrument and a summary of the reliability and validity results were provided. The statistics and the details varied by instrument but provided psychometric results given in the primary publication of the instrument. Given the difficulty of the task in evaluating some of the statistical methodology and descriptions of the psychometric properties, the breakout group was divided into 3 subgroups, and each subgroup

Table 4. Sleep instruments identified for further consideration based on a consensus of the scoring of the sleep domains, feasibility and psychometric properties.

Instrument	Truth[a]	Feasibility[b]	Domain of Sleep Identified (top 4)			
			Sleep Adequacy	Sleep Maintenance	Sleep Initiation	Daytime Functioning
Athens Insomnia Scale	13.6	2.3	x	x	x	x
Medical Outcome Study Sleep Measure	11	4.0	x	x	x	(x)[c]
Pittsburgh Sleep Diary	11	10.8	x	x	x	(x)
Women's Health Initiative Insomnia Rating Scale	8.8	5.5	x	x	x	

[a] Higher score indicated better psychometric properties. [b] Lower rank indicates better feasibility. [c] x indicates the domain is assessed; (x) indicates the domain is partially assessed.

reviewed 5 instruments. After their review, each subgroup provided opinions on the instruments they reviewed, and the breakout group chairs coordinated a discussion among all the breakout group participants and reached an accord on the scoring of the instruments (with a high score indicating good psychometric properties). The 4 highest scored instruments were: Athens Insomnia Scale, 13.6; Sleep Assessment Questionnaire, 13.0; Pittsburgh Sleep Diary, 12.0; MOS Sleep Measure, 11.0.

CONSENSUS BASED ON SLEEP DOMAINS, FEASIBILITY AND PSYCHOMETRIC PROPERTIES

Based on results from the 3 breakout groups (sleep domains considered most important, highest ranked sleep instruments on feasibility, and the highest scored sleep instruments on psychometric properties), the 4 sleep instruments identified for further consideration were the Athens Insomnia Scale, the MOS Sleep Measure, the Pittsburgh Sleep Diary and the Women's Health Insomnia Rating Scale (Table 4). Only the Athens Insomnia Scale assessed each of the top 4 sleep domains, and the MOS Sleep Measure and the Pittsburgh Sleep Diary assessed 3 of the domains and partially assessed the fourth domain of daytime functioning. Although the Pittsburgh Sleep Diary scored high on truth, it is difficult to complete and ranked low on feasibility. On the other hand, the Women's Health Insomnia Rating Scale was easy to complete and so ranked high on feasibility but did not score high on truth. Both the Athens Insomnia Scale and the MOS Sleep Measure scored high on truth and ranked high on feasibility.

In summary, sleep instruments have been evaluated on the domains assessed, feasibility, and psychometric properties. In terms of the OMERACT filter: truth relates to the domains assessed (content validity) and psychometric properties of validity and reliability; feasibility is directly related to administrative burden and applicability; and discrimination relates to the psychometric property of sensitivity or responsiveness.

A number of domains related to sleep have been reviewed, and several sleep instruments have been identified that may be applicable to RA patients, namely: Athens Insomnia Scale, the MOS Sleep Measure, the Pittsburgh Sleep Diary, and the Women's Health Insomnia Rating Scale. To further evaluate the sleep instruments identified, they should be considered in planned clinical trials of RA patients to assess their applicability. To further establish acceptability and applicability of the domains and the specific instruments, a Delphi exercise involving RA patients to further understand sleep quality from their perspective should be performed.

REFERENCES

1. Wells G, Anderson J, Boers M, et al. MCID/Low Disease Activity State Workshop. Summary, recommendations, and research agenda. J Rheumatol 2003;30:1115-8.
2. Kirwan J, Ahlmen M, de Wit M, et al. Incorporating the patient perspective into outcome assessment in rheumatoid arthritis — progress at OMERACT 7. J Rheumatol 2005;32:2250-56.
3. Asplund R. Sleep Disorders in the elderly. Drugs Aging 1999;14:91-103.
4. Moldofsky H. Sleep and pain. Sleep Med Rev 2001;5:385-96.
5. Lashley FR. A review of sleep in selected immune and autoimmune disorders. Holistic Nurs Pract 2003;17:65-80.
6. Smith MT, Wegener ST. Sleep disturbances in rheumatic diseases. Association of Rheumatic Health Professionals. Clinical Care in the Rheumatic Diseases, 3rd Ed. Ch 45, Atlanta: American College of Rheumatology; 2006:289-97.
7. Carr A, Hewlett S, Hughes R, et al. Rheumatology outcomes: The patient's perspective. J Rheumatol 2003;30:880-3.
8. Wolfe F, Michaud K, Li T. Sleep disturbance in patients with rheumatoid arthritis: Evaluation by Medical Outcomes Study (MOS) and visual analog sleep scales. J Rheumatol 2006;33:1942-51.
9. Abad VC, Sarinas PS, Guilleminault C. Sleep and rheumatologic disorders. Sleep Med Rev 2008;12:211-28.
10. Drewes AM, Svendsen L, Taagholt SJ, et al. Sleep in rheumatoid arthritis: a comparison with healthy subjects and studies of sleep/wake interactions. Br J Rheumatol 1998;37:71-81.
11. Association of Sleep Disorders Centers, Sleep Disorders Classification Committee, Roffwarg HP. Diagnostic classification of sleep and arousal disorders. Sleep 1979;2:1-137.
12. American Academy of Sleep Medicine. International classification of sleep disorders, 2nd ed. Diagnostic and coding manual. Westchester, IL: American Academy of Sleep Medicine; 2005.
13. Hays RD, Stewart AL. Sleep measures. In: Stewart AL, Ware JE, editors. Measuring functioning and well-being: The Medical Outcomes Study approach. Durham, NC: Duke University Press; 1992:235-59.

Appendix A. Reliability and validity of selected sleep instruments from the systematic review.

Instrument	Reliability	Validity
Athens Insomnia Scale	Internal Consistency: AIS was administered to 299 subjects (105 primary health care insomniacs, 144 psychiatric patients and 50 non-patient controls). Cronbach's alpha for AIS-8 was 0.89 and for AIS-5 was 0.87. Alpha remained unchanged when any individual item was removed. Mean item-total correlations were 0.67 for AIS-8 and 0.69 for AIS-5 ($p<0.001$) Factor Analysis: For both the AIS-8 and AIS-5, the entire scale emerged as a sole component (eigenvalues 4.56 and 3.29), with high percentages of variance explained (56.9% and 65.8%) and all items contributing almost equally Test-retest Reliability: For AIS-8 $r = 0.89$ and for AIS-5 $r = 0.88$. Individual items ranged from 0.77 to 0.86 ($p<0.001$)	Validity: In 299 subjects 176 insomniacs and 123 non-insomniacs were identified according to ICD-10 criteria. AIS-8 scores for the 2 groups were 11.05 (SD 4.89) and 2.28 (SD 2.56) ($p<0.001$). For a cutoff score of 6, sensitivity, specificity and correct case identification were 93%, 85% and 90%. For a cutoff score of 7, these values were 84%, 90% and 86%. Regression analysis showed 6 to be the optimum cutoff ($R^2 = .78$; beta = 0.77 ± 0.09; $p<0.001$). Positive and Negative Predictive Values varied between the general and psychiatric populations (PPV 41% and 86%, NPV 99% and 92%) Correlations of the AIS-8 and AIS-5 with the Jenkins Sleep Problems Scale were 0.90 and 0.85 ($p<0.001$)
Daily Sleep Diary	Reliability: 46 chronic pain patients were administered the DSD over 4 days. Coefficients of stability, for individual items ranged, from 0.38 to 0.62. All were statistically significant. Spearman Brown coefficients were also significant and ranged from 0.69 to 0.87. Repeated measures ANOVA showed that none of the DSD items changed across the recording period	Discriminant Validity: All individual items showed higher reliability coefficients than inter-item correlations supporting the ability of the individual measures to discriminate different aspects of sleep behavior Concurrent Validity: DSD items were correlated with retrospective summary measures of sleep from the Westhaven-Yale Multidimensional Pain Inventory. Duration of pain was related to delayed sleep onset ($r = 0.39$) and lower quality of sleep ($r = 0$-.34). Pain severity was related to fewer hours slept ($r = 0$ -.34) and delayed sleep onset ($r = 0$-.32). The DSD items also correlated with measures of both depression [Beck Depression (-0.40); Depression Adjective Checklist (-0.50)] and anxiety [State Trait Anxiety Inventory (-0.48)]
Dutch Sleep Disorders Questionnaire	125 sleep disorder patients and 20 normal university staff and students filled out the questionnaire Factor Analysis: 176 items from the Sleep Disorders Questionnaire (SDQ) were reduced by empirical methods and clinical judgment. 89 items were analyzed by principal components analysis (PCA) and varimax rotation. Items with factor loadings < 0.40 were eliminated. PCA yielded 26 components of which 4 were considered relevant: Insomnia (eigenvalue 13.46), narcolepsy (eigenvalue 5.10), sleep apnea (eigenvalue 8.35) and depression (eigenvalue 3.69) Cluster Analysis: K-means clustering resulted in 5 clusters: 1) Healthy, 2) Depression, 3) Insomnia, 4) Narcolepsy, 5) Apnea	Validity: Results of discriminant analysis matched the cluster analysis corroborating the validity of the cluster solution Polysomnographic diagnosis was available for 76 patients. Cluster analysis was most successful in narcolepsy and least successful in depression. Overall, 67% of subjects were classified correctly
Dysfunctional Beliefs and Attitudes About Sleep Questionnaire	Internal Consistency: In 145 older adults, 74 of whom were seeking treatment for insomnia, Cronbach's alpha for the total scale was 0.80 for good sleepers and 0.81 for poor sleepers. For the scale's 5 conceptually derived themes alphas for good sleepers, poor sleepers and total sample were as follows: 1) Consequences of Insomnia (0.79/0.69/0.77), 2) Control/Predictability of Sleep, (0.66/0.58/0.68), 3) Expectations (0.09/ -0.44/ -0.09), 4) Causal Attributions (0.24/0.37/0.31) and 5) Sleep Practices (0.58/0.56/0.56). Item-total correlations for the total sample ranged from 0.02 to 0.36, median 0.36	Validity: Significant between group differences for good and poor sleepers were found for Theme 1 (Hotelling's $T^2 = 0.115$ $p<0.03$, Theme 2 ($T^2 = .316$ $p< 0.1$, Theme 3 ($T^2 = 0.057$ $p< 0.05$, and Theme 5 ($T^2 = 0.423$ $p<0.0001$, but not for Theme 4

Epworth Sleepiness Scale	Test Retest Reliability: In 87 healthy medical students at 5 months, r=0.82 p<0.001. Internal Consistency: Cronbach's alpha was 0.88 for sleep disorder patients and 0.73 for students. Factor Analysis: 1 factor (possibly 2)	Validity: Administered to 150 patients with sleep disorders and 30 controls. Correlated with MSLT (Multiple Sleep Latency Test) sleep latency r= -0.514, N=27, p<0.01. Distinguished control subjects from obstructive sleep apnea, and narcolepsy or idiopathic hypersomnia
Insomnia Severity Index	Internal Consistency: In a population of 145 patients presenting with insomnia Cronbach's alpha = 0.74. Item total correlations varied from 0.36 to 0.67 with an average of 0.54 In a sample of 78 late life insomnia patients randomized to 4 treatment conditions internal reliability coefficients remained stable form 0.76 at baseline to 0.78 at followup	Validity: Correlation coefficients between individual ISI items and corresponding variables on a sleep diary for 145 patients presenting with insomnia were 0.35 for Sleep Onset Latency, 0.35 for Wake After Sleep Onset, and 0.35 for Early Morning Awakening. Correlation between the total ISI score and the diary Sleep Efficiency (ratio of Total Sleep Time to Total Time in Bed) was -0.19. All correlations were $p < 0.01$. In a sample of 78 late life insomnia patients randomized to four treatment conditions correlations for ISI variables and corresponding polysomnographic variables ranged from 0.07 to 0.45 at pretreatment and from 0.23 to 0.45 at post-treatment. Only the correlation for Sleep Onset Latency was statistically significant at pre treatment, whereas all correlations except Early Morning Awakening were significant post treatment. Correlations between the patient's and clinician's versions of the ISI at the two assessment periods were statistically significant as were correlations between the patient's and significant other's version
Leeds Sleep Evaluation Questionnaire	Reliability and consistency: In drug studies, dose related values were consistent, with linear relationships between dose level and self reported change on the Leeds SEQ	(Not reported in primary paper)
MOS Sleep	Internal Consistency: When tested on 2 samples (US general population and adults with neuropathic pain) internal consistency reliability estimates were acceptable for the scales: Sleep Disturbance (Cronbach's alpha = 0.80 and 0.82), Sleep Adequacy (alpha 0.82 and 0.76), Somnolence (alpha 0.63 and 0.73), Sleep Problems Index (alpha = 0.83, 0.78)	Validity: Neuropathic pain patients reported significantly more sleep disturbance and daytime somnolence as well as less quantity and adequacy of sleep than the general US population
Pittsburg Sleep Diary	Test Retest Reliability: In a sample of 96 healthy adults, over a 12 to 31 month delay, measures of both sleep timing and sleep quality showed correlations between 0.56 and 0.81 (p<0.001)	Validity: In a sample of 96 healthy adults, number and duration of awakenings as well as VAS ratings of sleep quality, mood and alertness showed statistically significant correlations with the PSQI. Agreement was shown between sleep diary and actiographic measures of both sleep time and quality
Pittsburg Sleep Quality Index	Test Retest Reliability: For global PSQI scores *r* was 0.85 ($p < 0.001$). Component scores within each subject group showed more variability across time but all these scores with the exception of medication use were significantly correlated ($r > 0.35$, $p < 0.05$) Internal Consistency: 7 component scores of PSQI had a high degree of internal consistency. Cronbach's $a = 0.83$. Mean component total *r* was 0.58	Validity: PSQI was administered to Controls, Depressives, Disorders of Initiating and Maintaining Sleep (DIMS) and Disorders of Excessive Somnolence (DOES). Diagnoses were based on clinical interviews, structured interviews and polysomnography. Control subjects differed significantly from all patient groups. DIMS and depressed patients had significantly higher scores than DOES patients Differences among all groups were further substantiated in MANCOVA testing for component scores across groups (Hotelling's $T^2 = 2.62$, $p < 0.001$). Distribution of global PSQI scores also differed between groups. Post hoc cutoff score of 5 correctly identified 88.5% (131/148) off all patients and controls (kappa = 0.75, $p < 0.001$, sensitivity 89.6% specificity 86.5%). Same cutoff correctly identified 84.4% DIMS, 88% DOES and 97% depressives. Group differences were also substantiated by polysomnographic results for sleep latency ($F = 4.53$, $p < 0.001$, sleep efficiency ($F = 5.78$, $p < 0.001$, sleep duration ($F = 4.82$, $p < 0.003$) and number of arousals ($F = 2.87$, $p < 0.04$). Group differences were not found for REM sleep or delta sleep. PQSI estimates of sleep variables were also compared to polysomnography. t tests showed no differences between PSQI estimates and lab findings for sleep latency but PSQI estimates for sleep efficiency and duration were greater then polysomnography ($t = 9.98$ and 4.50)

Sleep Assessment Questionnaire	Test-retest Reliability: 68 sleep disorder patients completed the questionnaire twice over 2 to 6 days. r=0.97. Factor Analysis: Principal component analysis with varimax rotation identified 5 factors – non restorative sleep, sleep schedule disorder, disturbed sleep, sleep apnea, and hypersomnolence. Cronbach's alpha did not increase when individual items were removed from each of the factors indicating that the questions were homogeneous	Criterion Validity: Subjects were asked 2 questions: Do you have trouble sleeping? (QA) and Do you have trouble staying awake? (QB). QA correlated with factors non-restorative sleep (r=0.67, p<0.0001) and disturbed sleep (r=0.63, p<0.0001) and QB correlated with hypersomnolence (r=0.49, p<0.0001). Discriminant Validity: 289 sleep patients had higher mean total SAQ than 30 controls (26.0 ± 8.6 vs 10.8 ± 5.7 p<0.0001). The SAQ showed favourable sensitivity and specificity for discriminating patients from normal subjects Of the patient group, those with Sleep Apnea and Periodic Limb Movements had means on all SAQ factors with the exception of Sleep Schedule Disorder that were significantly different from normals
Sleep Disorders Questionnaire	Test-retest Reliability - Full Scale: Over 2 weeks, in 71 subjects without sleep complaints, item reliabilities ranged from r=0.999 to 0.163 (all except 3 p< 0.0001). Mean r = 0.704. Completion rate was 95.7% In 130 sleep disordered patients over 3-4 months, correlations ranged from r=0.308 to 0.985 (all p<0.0001). Mean r = 0.636. 28 items achieved r>0.80 Canonical Discriminant Functional Analysis: Four Scales were produced: Sleep apnea (SA), Narcolepsy (NAR), Psychiatric Sleep Disorder (PSY), Periodic Limb Movement (PLM) Intercorrelations of the 4 scales were SA-NAR = 0.14, SA-PSY = - 0.20, SA-PLM = 0.34, NAR-PSY = 0.27, NAR-PLM = 0.38, PSY-PLM = 0.48 Test-retest Reliability - Subscales: In 130 sleep-disorder patients over 4 months Spearman *rho* was as follows: SA 0.842, NAR 0.753, PSY 0.848, and PLM 0.817. Internal Consistency - Subscales: Cronbach's alpha was as follows: SA 0.855, PLM 0.695, PSY 0.800 NAR 0.853	Internal Validity - Subscales: For all subscales except PLM the "characteristic" scale for a patient group had a significantly higher mean in that group compared to all other groups. Sensitivity - Subscales: SA - Male 0.85, Female 0.88 NAR - Male 0.84, Female 0.80 PSY - Male 0.79, Female 0.79 PLM - Male 0.67, Female 0.65 Specificity - Subscales: SA - Male 0.76, Female 0.81 NAR - Male 0.68, Female 0.72 PSY - Male 0.65, Female 0.64 PLM - Male 0.46, Female 0.49
Sleep Dissatisfaction Questionnaire	Factor Analysis: 8 factors were found, of which 6 were interpretable: After Effects (AE), Sleep Attitudes (SA), Mental Activity (MA), Sleep Maintenance (SM), Dissatisfaction (D), and Sleep Onset (SO)	Validity: Correlation of factors with the following variables was statistically significant. AE: worry 0.22, neuroticism 0.29, sleep aspiration 0.18, problem nights per week -0.20 SA: sleep aspiration -0.20, problem nights per week 0.28, nights without sleep onset problem 0.39 MA: worry 0.25, neuroticism 0.44, age -0.32, problem nights per week 0.26 SM: worry 0.25, age 0.26, sleep aspiration -0.20, nights with sleep onset problem -0.27, nights without sleep onset problem -0.26, frequency of mid-sleep awakenings 0.48, time to return to sleep 0.41 D: age 0.21, sleep problem 0.22, morning awakenings per week 0.22 SO: sleep duration-0.26, problem nights per week nights 0.53, nights with sleep onset problem 0.33, nights without sleep onset problem 0.42
Stanford Sleepiness Scale	(Not reported in primary paper)	Validity: 5 healthy males were tested 4x/day (2x for Wilkinson Addition Test plus a memory test and 2x for Wilkinson Vigilance Test) for 6 consecutive days. They rated themselves every 15 minutes during a 16 hour day on the SSS. On every night, except night 4, they maintained a standard bedtime, with 8 hours in bed. On night 4 they underwent all night sleep deprivation. Mean SSS ratings correlated non-significantly with performance on the Wilkinson Addition Test (r=0.67, SD 0.34) and Wilkinson Vigilance Test (r=.70, SD .31). Performance on these tests was found to decrease as SSS ratings increased by a mean of 2.91, SD = 1.67. Discrete SSS ratings correlated r=.47 with performance on the memory test

Women's Health Initiative Insomnia Rating Scale	The population consisted of 66,269 women participating in the Women's Health Initiative (WHI). A random sub-sampling technique was used. Factor Analysis: The WHIIRS had a stable one-factor solution and multigroup structural equation modeling revealed measurement invariance across age and race-ethnic groups. Test Retest Reliability: 0. 96 for same day administration and 0.66 after one year. Internal Consistency: Mean alpha was 78. 89% of the samples had reliability coefficients > 0.75	Construct Validity: All trends and correlations between the WHIIRS and measures of related constructs were in the expected direction. The WHIIRS showed an almost zero correlation with the Negative Emotional Expressiveness scale with which it was predicted to have no relationship. In 459 women who had reported abnormal sleep duration, the mean WHIIRS score indicated more sleeping difficulty [7.62 (SD 4.92)] as compared with the normative sample [6.61 (SD 4.45)] In the same 459 women, who underwent wrist actiography, 2 groups were formed by applying the definition of insomnia as < 85% efficacy and a latency of < 30. The mean WHIIRS score in the insomnia group was 9.08 (SD = 5.58, n = 100) and the mean in the other group was 6.76 (SD = 4.53, n=322). Mean difference was 0.48 pooled SD ($p<.001$) suggesting approximately .05 SD as being a clinically meaningful difference. Trend analysis indicated a significant linear relationship between the WHIIRS and Waking After Sleep Onset (WASO) as measured by actiograph reflecting an increase in scale scores as WASO increased. In addition, the correlations between the actiograph measures WASO, Latency, Efficiency and Duration showed that these objective variables correlated most highly with the WHIIRS items intended to tap into the same insomnia construct. The correlation between the total WHIIRS score and Efficiency was negative (-2.00 $p<.05$) as would be expected. WHIIRS correctly predicted those with and without insomnia with a probability of .65. Using a cutoff score of 9 the sensitivity, specificity, positive predictive value and negative predictive value were 0.53, 0.67, 0.23, and 0.88 respectively

14. Subar A, Ziegler R, Thompson F, et al, for the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. Is shorter always better? Relative importance of questionnaire length and cognitive ease on response rates and data quality for two dietary questionnaires. Am J Epidemiol 2001;153:404-9.
15. Gillberg M, Kecklund G, Akerstedt T. Relations between performance and subjective ratings of sleepiness during a night awake. Sleep 1994;17:236-41.
16. Soldatos CR, Dikeos DG, Paparrigopoulos TJ. Athens Insomnia Scale: Validation of an instrument based on ICD-10 criteria. J Psychosom Res 2000;48:555-60.
17. Partinen M, Gislason T. Basic Nordic Sleep Questionnaire (BNSQ): A quantitated measure of subjective sleep complaints. J Sleep Res 1995;4 Suppl 1:150-5.
18. Cote KA, Ogilvie RD. The Brock Sleep and Insomnia Questionnaire: Phase 1. Sleep Res 1993;22:356.
19. Haythornthwaite JA, Hegel MT, Kerns RD. Development of a sleep diary for chronic pain patients. J Pain Symptom Manage 1991;6:65-72.
20. Sweere Y, Kerkhof GA, De Weerd AW, Kamphuisen HA, Kemp B, Schimsheimer RJ. The validity of the Dutch Sleep Disorders Questionnaire (SDQ). J Psychosom Res 1998;45:549-55.
21. Morin CM, Stone J, Trinkle D, Mercer J, Remsberg S. Dysfunctional beliefs and attitudes about sleep among older with and without insomnia complaints. Psychol Aging 1993;8:463-7.
22. Johns MW. A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. Sleep 1991;14:540-5.
23. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep 1992;15:376-81.
24. Espie CA, Brooks DN. An evaluation of tailored psychological treatment of insomnia. J Behav Ther Exp Psych 1989;20:143-53.
25. Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. Sleep 1997;20:835-43.
26. Harvey KJ, Espie CA. Development and preliminary validation of the Glasgow Content of Thoughts Inventory (GCTI): A new measure for the assessment of pre-sleep cognitive activity. Br J Clin Psych 2004;43:409-20.
27. Pavlova M, Berg O, Gleason R, Walker F, Roberts S, Regestein Q. Self-reported hyperarousal traits among insomnia patients. J Psychosom Res 2001;51:435-41.
28. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001;2:297-307.
29. Jenkins CD, Stanton B-A, Niemcrym SJ, Rose RM. A scale for the estimation of sleep problems in clinical research. J Clin Epidemiol 1988;41:313-21.
30. Keklund G, Akerstedt T. Objective components of individual differences in subjective sleep quality. J Sleep Res 1997;6:217-20.
31. Akerstedt T, Gillberg M. Subjective and Objective Sleepiness in the Active Individual. Int J Neurosci 1990;52:29-37.
32. Parrott AC, Hindmarch I. The Leeds Sleep Evaluation Questionnaire in psychopharmacological investigations — a review. Psychopharmacology 1980;71:173-9.
33. Hays RD, Martin SA, Sesti AM, Spritzer KL. Psychometric properties of the Medical Outcomes Study Sleep Measure. Sleep Med 2005;6:41-4.
34. Maldonado CC, Bentley AJ, Mitchell D. A pictorial sleepiness scale based on cartoon faces. 2004;27:541-8.
35. Monk TH, Reynolds CF, Kupfer DJ, et al. The Pittsburgh Sleep Diary. Sleep Res 1994;3:111-20.
36. 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;28:193-213.
37. Webb WB, Bonnet M, Blume G. A post-sleep inventory. Percept Motor skills 1976;43:987-93.
38. Nicassio PM, 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;23:263-71.
39. Pires De Souza JC. Quality of life and insomnia in university psychology students. Human Psychopharmacology 1996;11:169-84.
40. Rombaut N, Maillard F, Kelly F, Hindmarch I. The Quality of Life of Insomniacs Questionnaire (QOLI). Med Sci Res 1990;18:845-7.
41. Violani C, Lucidi F, Robusto E, Devoto A, Zucconi M, Ferini Strambi L. The assessment of daytime sleep propensity: a comparison between the Epworth Sleepiness Scale and a newly

developed Resistance to Sleepiness Scale. Clin Neurophysiol 2003;114:1027-33.
42. Van Knippenberg FC, Passchier J, Heystech D, et al. The Rotterdam Daytime Sleepiness Scale: a new daytime sleepiness scale. Psychol Rep 1995;76:83-7.
43. Spoormaker VI, Verbeek I, van den Bout J, Klip EC. Initial Validation of the SLEEP-50 Questionnaire. Behav Sleep Med 2005;3:227-46.
44. Cesta A, Moldofsky H, Sammut C. The University of Toronto Sleep Assessment Questionnaire (SAQ). Sleep Res 1996;25:486.
45. Cesta A, Moldofsky H, Sammut C. The University of Toronto Sleep Assessment Questionnaire (SAQ). Sleep Res 1997;26:646.
46. Kazarian SS, Howe MG, Csapo KG. Development of the Sleep Behavior Self-Rating Scale. Behav Ther 1979;10:412-7.
47. Adan A, Fabbri M, Vincenzo N, Prat G. Sleep Beliefs Scale (SBS) and circadian typology. J Sleep Res 2006;15:125-32.
48. Douglass AB, Bornstein R, Nino-Murcia G, et al. The Sleep Disorders Questionnaire I: creation and multivariate structure of SDQ. Sleep 1994;17:160-7.
49. Coyle K, Watts FN. The factorial structure of sleep dissatisfaction. Behav Res Ther 1991;29:513-20.
50. Zammit GK. Subjective ratings of the characteristics and sequelae of good and poor sleep in normals. J Clin Psych 1988;44:123-30.
51. Parrott AC, Hindmarch I. Factor analysis of a sleep evaluation questionnaire. Psychol Med 1978;8:325-9.
52. Mastin DF, Bryson J, Corwyn R. Assessment of sleep hygiene using the Sleep Hygiene Index. J Behav Med 2006;29:223-7.
53. Neitzert Semler C, Harvey AG. Monitoring for sleep-related threat: a pilot study of the Sleep Associated Monitoring Index (SAMI). Psychosom Med 2004;66:242-50.
54. Yi H, Shin K, Shin C. Development of the Sleep Quality Scale. J Sleep Res 2006;15:309-16.
55. Domino G, Blair G, Bridges A. Subjective Assessment of Sleep by Sleep Questionnaire. Percept Mot Skills 1984;59:163-170.
56. Lacks P. Behavioral Treatment of Persistent Insomnia. Elmsford, New York: Pergamon Press;1987.
57. Monk TH, Buysse DJ, Kennedy KS, Potts JM, DeGrazia JM, Miewald JM. Measuring Sleep Habits Without Using a Diary: The Sleep Timing Questionnaire. Sleep 2003;26:208-12.
58. Rosenthal L, Roehrs TA, Roth T. The Sleep-Wake Activity Inventory: A self-report measure of daytime sleepiness. Biol Psychiatry 1993;34:810-20.
59. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC, Quantification of sleepiness: a new approach. Psychophysiology 1973;10:431-6.
60. Ellis BW, Johns MW, Lancaster R, Raptopoulos P, Angelopoulos N, Priest RG. The St. Mary's Hospital Sleep Questionnaire: A study of reliability. Sleep 1981;4:93-7.
61. Leigh TJ, Bird HA, Hindmarch I, Constable PD, Wright, V. Factor analysis of the St. Mary's Hospital Sleep Questionnaire. Sleep 1988;11:448-53.
62. Snyder-Halpern R, Verran JA. Instrumentation to describe subjective sleep characteristics in healthy subjects. Res Nurs Health 1987;10:155-63.
63. Levine DW, Kripke DF, Kaplan RM, et al. Reliability and validity of the Women's Health Initiative Insomnia Rating Scale. Psychol Assess 2003;15:137-48.
64. 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. Psychol Assess 2004;15:123-36.