

SCIENTIFIC INVESTIGATIONS

Psychometric Evaluation of the Insomnia Symptom Questionnaire: a Self-report Measure to Identify Chronic Insomnia

Michele L. Okun, Ph.D.¹; Howard M. Kravitz, D.O., M.P.H.²; Mary Fran Sowers, Ph.D.³; Douglas E. Moul, M.D.¹; Daniel J. Buysse, M.D.¹; Martica Hall, Ph.D.¹

¹University of Pittsburgh, Department of Psychiatry, Pittsburgh, PA; ²Rush University Medical Center, Department of Psychiatry and Department of Preventive Medicine, Chicago, IL; ³University of Michigan Center for Integrated Approaches to Complex Diseases, Ann Arbor, MI

Study Objectives: The objective was to psychometrically evaluate the Insomnia Symptom Questionnaire (ISQ), a self-report instrument designed to establish a clinically relevant case definition of insomnia consistent with widely used insomnia classification criteria, using methods from classical test theory and item response theory (IRT).

Methods: The ISQ was evaluated using IRT algorithms in a cohort of 362 pre-, peri- and post-menopausal women recruited for the SWAN (Study of Women's Health Across the Nation) Sleep Study. This yielded a dichotomous outcome consistent with the presence/absence of insomnia. The internal consistency and criterion validity of the dichotomized ISQ were compared to traditional measures of sleep from sleep diaries, polysomnography, and the Pittsburgh Sleep Quality Index using kappa statistics, and indices of sensitivity, specificity, positive and negative predictive value (PPV), and likelihood ratio tests (LRs).

Results: The ISQ identified 9.8% of the sample as meeting insomnia, consistent with established diagnostic criteria. Reliability was established with Cronbach α ($\alpha = 0.89$). The ISQ had high specificity

(> 90%), but sensitivity, PPV, NPV, and LR_s varied according to which sleep measure was used. Concurrent validity was not confirmed with any of the traditional sleep summary measures (κ 's < 0.30).

Conclusions: The ISQ captures the multidimensionality of insomnia better than traditional sleep measures as it ascertains symptoms of insomnia that are based on DSM-IV and RDC criteria. The high specificities suggest that the ISQ has a high probability of correctly identifying those without insomnia and would be a cost-effective tool in large observational studies in which the prevalence of insomnia is likely to be about 10%. Further evaluation of the ISQ, including validation against clinical interviews, is warranted.

Keywords: Insomnia; questionnaire; sleep; menopausal transition; validity

Citation: Okun ML; Kravitz HM; Sowers MF; Moul DE; Buysse DJ; Hall M. Psychometric evaluation of the insomnia symptom questionnaire: a self-report measure to identify chronic insomnia. *J Clin Sleep Med* 2009;5(1):41-51.

Insomnia is a major public health problem affecting millions of individuals, along with their families and communities.¹ Insomnia is defined as difficulty initiating or maintaining sleep, or unrefreshing sleep despite adequate opportunity for sleep. It is the most commonly reported sleep disturbance, affecting up to 30% of the adult population.²⁻⁴ Women are 1.4 times as likely as men to complain of insomnia symptoms,⁵ and this number significantly increases with the transition into the menopausal period.^{3,6-8} Recent epidemiological data indicate that approximately 40% of women between the ages of 40 and 55 years report recent sleep difficulties that resemble symptoms of insomnia.⁶

Diagnostic criteria for insomnia from the American Academy of Sleep Medicine and the American Psychiatric Association include: (1) a complaint of problems initiating or maintaining sleep, and/or that sleep is nonrestorative or poor in quality; (2) the duration of the sleep complaint is at least 1 month; and (3) complaints of significant impairment in social, occupational

or daily functioning.^{9,10} Although these criteria do not include frequency of occurrence, a frequency criterion of ≥ 3 nights per week is commonly used.¹¹ Clinical interviews conducted by trained clinicians are used to establish a clinical diagnosis of insomnia, but the use of a fully structured interview¹² is often impractical in large epidemiological studies because it is time-consuming and costly.^{13,14} Cost-effective metrics for assessing insomnia are needed.

Various symptom severity questionnaires have incorporated diagnostic criteria for chronic insomnia.¹⁵ Several of these correspond directly to specific established diagnostic criteria by assessing sleep complaints, duration and/or daytime consequences. These include the Insomnia Severity Index (ISI),¹⁶ the Medical Outcomes Study (MOS) Sleep measure,¹⁷ the Women's Health Initiative Insomnia Rating Scale (WHIIRS),¹⁸ the Athens Insomnia Scale,¹⁹ and the Short Insomnia Questionnaire (SDQ).²⁰ However, none of these instruments establishes a case definition of insomnia that would be an appropriate analog of clinical interviews for epidemiological studies, and none provides a dichotomous outcome based on a case definition that would allow a clinician or researcher to confidently categorize participants.¹⁵

Other sleep instruments assess specific sleep characteristics or quality, but were not designed to identify insomnia per se. For instance, the Pittsburgh Sleep Quality Index (PSQI)²¹ is a

Submitted for publication February, 2008

Submitted in final revised form August, 2008

Accepted for publication August, 2008

Address correspondence to: Martica Hall, Ph.D. University of Pittsburgh, WPIC, 3811 O'Hara St E1119, Pittsburgh, PA 15213; E-mail: hallmh@upmc.edu

questionnaire with established reliability and validity that is frequently used for measuring sleep quality and identifying good and bad sleepers. However, it was not designed to assess insomnia based on diagnostic criteria. The PSQI's reliability and validity for identifying people who have difficulty initiating or maintaining sleep, a necessary symptom of insomnia, may be improved by using a more stringent cutoff score (> 10).²⁰⁻²² Nevertheless, the PSQI does not ascertain frequency of sleep symptoms nor specific information on daytime consequences.

Based on the above, we identified the need for a self-report questionnaire guided by established diagnostic criteria that could be used to identify insomnia in large epidemiological or health services research studies. We administered the 13-item Insomnia Symptom Questionnaire (ISQ) (Appendix 1) to pre-, peri-, and post-menopausal women in a large observational study and compared the reports of symptoms from this instrument to other measures of sleep including the PSQI, sleep diaries, and nocturnal polysomnography (PSG) data. There was no true "gold standard" for diagnosing insomnia in our sample as we had no clinical interview. Rather, the available sleep measures were used to suggest the presence of a sleep disorder. Therefore, the analyses focused on the internal consistency and criterion (i.e., concurrent) validity of the ISQ. The aims of the present analyses were (1) to describe the development of a scoring algorithm using item response theory (IRT) that yielded a clinically relevant case definition of insomnia, and (2) to present the psychometric properties of the ISQ.

METHOD

Participants

Participants were 370 mid-life women who took part in a cross-sectional study of sleep during the menopausal transition. They were recruited from 4 (Chicago, IL; Detroit area, MI; Oakland, CA; and Pittsburgh, PA) of 7 study sites participating in the Study of Women's Health Across the Nation (SWAN), a longitudinal study of menopause and aging. The recruitment methods and study design for the Core SWAN Study have been described.²³ All Core SWAN participants participating in the year 5, 6, or 7 annual follow-up visits (2003–2005) were approached for SWAN Sleep Study participation. Exclusions for Sleep Study participation were hormone therapy (HT) use; active chemotherapy or radiation; oral corticosteroid use; regular participation in shiftwork; regular consumption of > 4 alcoholic drinks/day; and noncompliance with Core SWAN study procedures (missed $> 50\%$ of annual visits, refused annual visit blood draw). During the final year of the sleep study, postmenopausal women were also included to provide evaluations of sleep across the full spectrum of menopausal transition stages. Eight women had incomplete or missing data for the ISQ, leaving an analyzable sample of 362 women. Informed consent was obtained in accordance with institutional review board-approved protocols at each SWAN site.

Procedure

Sleep data collection included PSG, daily sleep diaries, and self-report sleep symptom questionnaires (ISQ, PSQI). PSG

sleep studies were conducted in participants' homes during the first 3 days of the protocol which spanned a full menstrual cycle or up to 35 days, whichever was shorter. Participants were scheduled for the SWAN Sleep Study protocol based upon the timing of their menstrual cycle. Each woman began the sleep protocol at the initiation of menstrual bleeding (between days 3-6 of the early follicular phase) to allow greater comparability in considering cycle effects on sleep. If the participant had been without menses or was not regularly cycling for at least 4 months she was considered peri- or late-peri-menopausal and the study was scheduled at her convenience. Throughout the protocol, participants completed daily sleep-wake diaries at bedtime and upon awakening. The ISQ and the PSQI data used in this report were completed on Day 4 of the protocol.

Measures

The Insomnia Symptom Questionnaire (ISQ) (Appendix 1) is a 13-item self-report instrument designed to identify insomnia. ISQ questions were based on the American Psychiatric Association's fourth edition of the Diagnostic and Statistical Manual (DSM-IV) criteria for primary insomnia¹⁰ and are consistent with the American Academy of Sleep Medicine's (AASM) Research Diagnostic Criteria (RDC).²⁴ However, the ISQ does not assess whether the sleep disturbance occurs exclusive of another sleep disorder, mental disorder, or substance or general medical condition (criteria C-E in DSM-IV for primary insomnia); nor does it ask if the participant has adequate opportunity to sleep (RDC criteria). Questions on the ISQ follow a stepped approach: (1) ascertaining the presence of a complaint of difficulty initiating or maintaining sleep, or feeling that the sleep was nonrestorative or unrefreshing; (2) determining the frequency of complaints (questions 1-5) and duration of these symptoms; and (3) evaluating the severity of daytime correlates of the sleep complaint(s) (questions 6-13). The ISQ items contain multiple choices on an ordinal scale to assess the presence, frequency and/or severity of the complaint. For example, the "During the past month have you had difficulty falling asleep?" item includes choices ranging from 0 = never to 5 = always (5-7 days per week), and a follow-up contingency question asks about the problem's duration with open-ended responses for weeks, months, or years. Similarly, questions 6-13 assess the extent to which the individual's endorsed sleep complaints affect daytime activities, with response choices ranging from 0 = not at all to 4 = extremely.

*The Pittsburgh Sleep Quality Index (PSQI)*²¹ is a 19-item questionnaire used to measure sleep quality complaints. Seven component scores assess habitual duration of sleep, nocturnal sleep disturbances, sleep latency, sleep quality, daytime dysfunction, sleep medication usage, and sleep efficiency. The 7 components (range 0-3) are summed to yield a measure of global sleep quality with a range of 0 (good sleep quality) to 21 (poor sleep quality). The PSQI and its psychometric properties have been characterized in various populations, including patients with insomnia and depression and in healthy good sleeper controls.^{21,22} These reports have documented adequate sensitivity and specificity using a cutoff score > 5 to identify those with poor sleep quality in depressed and normal sleepers, and a cutoff score > 10 to identify poor sleep quality among people with insomnia.

Self-reported and PSG assessed sleep variables. Three pathognomic symptoms of insomnia are difficulty falling asleep, difficulty staying asleep, and unrefreshing sleep. Sleep onset latency, wake after sleep onset, and sleep efficiency are 3 derived sleep variables used to quantify the first 2 of these symptoms. Unrefreshing sleep is defined from a subjective perspective. Sleep onset latency (SOL) is defined as the time one “time tried to go to sleep” to sleep onset. Wake after sleep onset (WASO) is defined as the total number of minutes spent awake following sleep onset. Time in bed (TIB) is defined as the total amount of time from reported “time tried to go to sleep” to the time of the reported final awakening from sleep. Total sleep time (TST) is defined as minutes of time spent asleep. Sleep efficiency (SE) is defined as the ratio of TST to TIB multiplied by 100.

*Sleep diaries*²⁵ were used to capture daily measures of self-reported sleep parameters at both a morning and evening recording time. Variables included mean values for SOL, WASO, TST, TIB, and SE calculated from all available sleep diary data.

Ambulatory polysomnography (PSG) studies were conducted in participants’ homes. Recordings included bilateral central referential EEG (C_3 and C_4 , referenced to $A_1 + A_2$), bilateral electro-oculograms (EOG), submental is electromyogram (EMG), a modified V2 lead electrocardiogram (EKG), and inductance plethysmography abdominal and thoracic belts. On Night 1, additional data (nasal pressure monitoring, oral-nasal thermistors, and fingertip oximetry, and bilateral anterior tibialis EMG leads) were collected to assess sleep disordered breathing and periodic leg movements. Sleep was visually scored in 20-s epochs using standard scoring criteria,²⁶ supplemented by apnea-hypopnea criteria derived from American Academy of Sleep Medicine recommendations,²⁷ and standard rules for scoring periodic limb movements.²⁸ Summary PSG variables (averaged over Nights 2 and 3) included SOL, WASO, TST, TIB, and SE.

Psychometric and Statistical Analyses

Item response theory (IRT) was used to determine the optimal cutoffs for individual questions and yield a dichotomous outcome (insomnia or no insomnia) that most closely reflects established diagnostic criteria for insomnia. Conceptual guidelines and established psychometric techniques, including internal consistency and criterion validity measures, were used to evaluate validity between the dichotomized ISQ and traditional sleep measures. In order to restrict our analyses to individuals without clinically significant sleep disordered breathing, women with an AHI ≥ 15 were excluded from all analyses. The final sample for the psychometric analyses included 266 women, 73.5% of the initial cohort.

ITEM ANALYSIS AND DEVELOPMENT OF SCORING ALGORITHM

In IRT, logistic regression equations are used to model a person’s odds of a response choice from questionnaire items based upon that respondent’s location along a latent dimension of an ability or trait (here insomnia). This modeling locates a response along the overall dimension based upon the respondent’s total scale-score and, based upon this estimate, it also approximated each item’s location (“item threshold”) and relevance (“discrimination”) of reporting along a severity dimension “theta.”²⁹

Simultaneously, the standard error of location of a response on the latent trait was estimated.³⁰ The collection of item parameters comprised the representation of the overall severity of the latent dimension of insomnia. We used a 2-parameter model. The first parameter represented discrimination, namely, how “decisive” the question was, compared to other questions. The second parameter represented thresholds, namely where the item information was located along the spectrum severity.

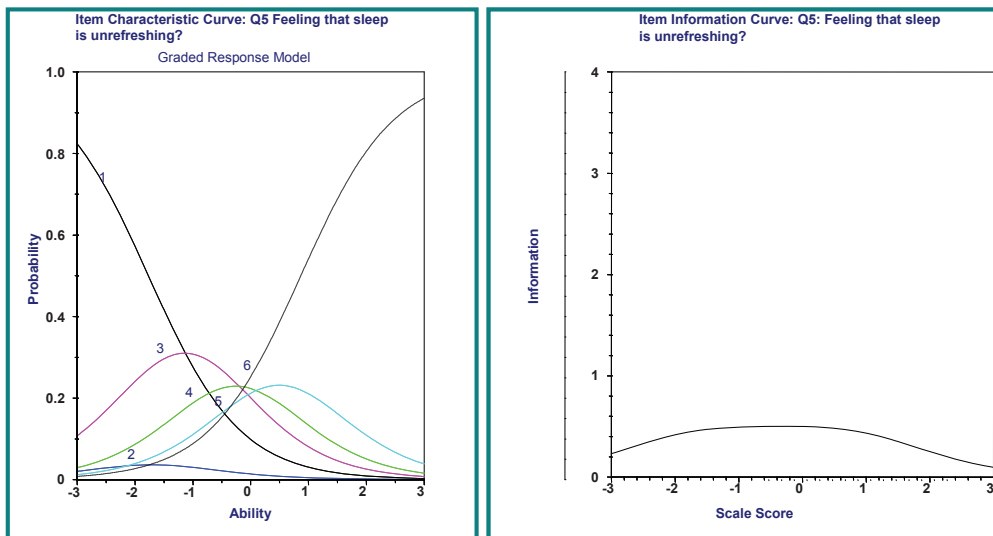
IRT was also used to identify the questions that uniquely contributed the most useful information to an ISQ insomnia definition (item-category reduction). Item-category reduction assists in reducing redundancy of questions. It is also desirable for the questionnaire to be unidimensional. We used scree plot and eigenvalues when modeling across all item categories (with the item being considered a “testlet”) conjointly estimating every item information curve (IIC) in the collection of items. The information curve data then allow one to justify where the number of item-categories could be reduced and compared the relative information provided by each item. We began the item-category reduction process by heuristically evaluating the item-category probability curves or item characteristic curves (ICC) for each item, which shows the predicted logistic function probability of responding in the item choice category, given where the respondent is estimated to be on the trait spectrum.³¹ This process was completed iteratively until a dichotomous outcome could be determined for each item and for the entire ISQ.

STATISTICAL ANALYSES

Descriptive statistics were used to characterize the subjective (questionnaires and diary) and PSG sleep measures for the sample. Traditional psychometric characteristics including content and face validity, internal validity, and criterion (concurrent) validity were then examined. Internal reliability was characterized using Cronbach α statistic.³² In the absence of a gold standard, concurrent validity and indicators of diagnostic accuracy (e.g., sensitivity, specificity, and predictive value) were estimated by comparing the dichotomized ISQ to widely used indicators of insomnia (e.g., PSG- and diary-assessed indicators of difficulty falling and staying asleep).³³ Lastly, likelihood ratios (LR) were used to evaluate how adequately the ISQ performed as a diagnostic test. This approach has the advantage of being less likely to change with the prevalence of the disorder in different samples as compared to estimates of sensitivity and specificity.³⁴ “LR positive” indicates how much more likely it would be to find a positive ISQ in those with insomnia compared to those without insomnia. Generally, LR values > 10 suggest a strongly positive result, ≥ 5 -10 indicate a moderate result, ≥ 2 -5 is a small result, and LR values of 1-2 suggest the questionnaire is non-informative.³⁴

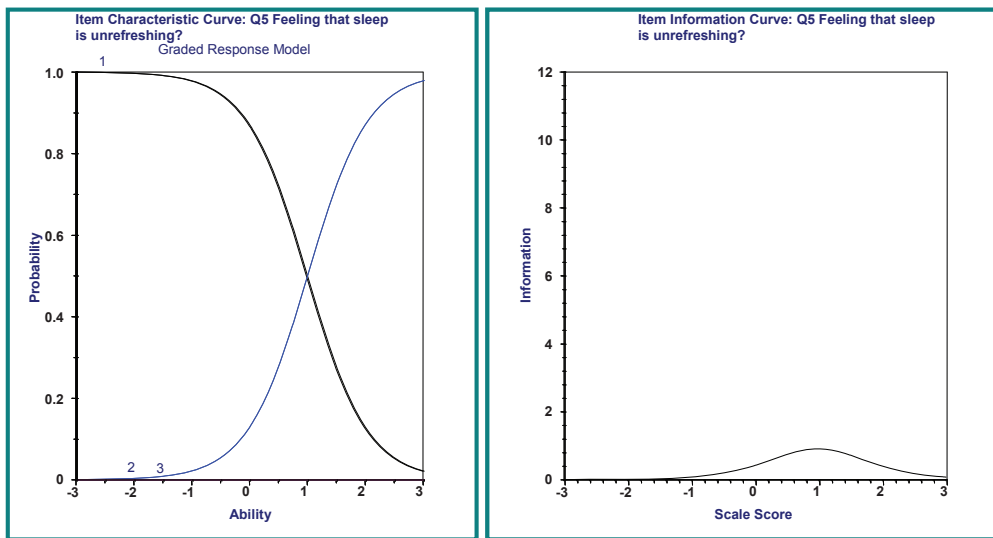
Insomnia estimates used to evaluate the validity of the ISQ were derived from three sources: PSQI, sleep diaries, and PSG. We dichotomized the PSQI global score at 2 cutoffs (> 5 and > 10) reflecting the values used in previous mixed-subjects samples²¹ and samples of insomnia subjects.²⁰⁻²² The decision to use comparison values for SOL (≥ 31 minutes), WASO (≥ 31 minutes), and SE ($< 85\%$) from the sleep diaries and PSG was based on various reports and recommendations for quantitative insomnia criteria.^{24,11,35} For the sleep diary data, the frequency of the sleep complaint was also assessed in accordance with the RDC.²⁴

Figure 1a



Category legends		
1 = Not at all	4 = Sometimes (1-2X/Week)	
2 = Don't Know	5 = Frequently (3-4X/Week)	
3 = Rarely (<1X/week)	6 = Always (5-7X/Week)	

Figure 1b



Category legends		
1 = < 3x/Week	2 = 3+X/Week	3 = Missing

Figure 1—Item Characteristic Curve (ICC) and Item Information Curve (IIC) Illustration. These graphs display parent polytomous and child dichotomous item characteristic curves obtained as a part of deriving dichotomous cutpoints for the ISQ questions. For each ICC curve, the inflection point represents the location of item thresholds along the spectrum of responding (“Theta”) where near-neighbor responders were 50% likely to endorse a more severe or less severe item category. The IIC curve depicts the information location and strength provided by the item, compared to other questionnaire items, along the same spectrum, based upon the information provided by each ICC curve. In the ICC graphs for the polytomous responses concerning the frequency of difficulty falling asleep, it can be appreciated that the item thresholds are comparatively aggregated in the region of Theta = +1. In the corresponding IIC graphs, the information provided by the summation of the item categories can be noted. The IIC item evaluation points to where the item provides its overall information. In the dichotomous simplification of this item, where the responses are dichotomized, the binomial IIC is located in the approximate location where the thresholds are located in the parent polytomous item. The IIC for the binomial child item illustrates that the dichotomization of the parent polytomous item has preserved the information profile of the parent item in the approximate region where the case:control separation could be plausibly drawn. The decision of cutpoint was made after several data runs, as the number of polytomous categories was sequentially reduced.

Table 1—Subjective and Polysomnographic Sleep Characteristics for Women Who Meet the Case Definition (Insomnia) and Women Who do not Meet the Case Definition (No Insomnia)

	No Insomnia N = 320 (89.4%)	Insomnia N = 38 (10.6%)	t-test
PSQI global score	6.4 ± 2.3	9.5 ± 3.2	t = 16.6, df = 345, p < 0.001
Diary Data			
Sleep Onset Latency (minutes)	14.6 ± 10.7	24.9 ± 18.6	t = -3.4, df = 352, p = 0.002
Wake after Sleep Onset (minutes)	14.0 ± 13.8	24.5 ± 16.5	t = -4.4, df = 352, p < 0.001
Total Sleep Time (minutes)	401.8 ± 47.9	379.1 ± 77.4	t = 1.8, df = 352, p > 0.05
Time in Bed (minutes)	492.7 ± 70.1	518.1 ± 118.4	t = -1.3, 352, p > 0.05
Sleep Efficiency (%)	93.2 ± 4.6	87.5 ± 7.2	t = 4.7, df = 352, p < 0.001
Polysomnography Data			
AHI (events per hour)	9.6 ± 13.4	16.6 ± 23.2	t = -1.8, df = 356, p > 0.05
Sleep Onset Latency (minutes)	21.3 ± 20.1	24.6 ± 26.7	t = -0.94, df = 356, p > 0.05
Wake after Sleep Onset (minutes)	52.0 ± 30.2	67.9 ± 43.2	t = 8.5, df = 356, p = 0.03
Total Sleep Time (minutes)	376.0 ± 53.6	351.5 ± 69.8	t = 3.6, df = 356, p < 0.05
Time in Bed (minutes)	449.0 ± 58.8	444.1 ± 65.0	t = 0.50, df = 356, p > 0.05
Sleep Efficiency (%)	83.8 ± 7.7	79.1 ± 10.2	t = 7.1, df = 356, p = 0.009

Table 2a—Cross-Tabulations Between PSQI Global Score (Cut-off of >5 and >10) and ISQ Classifications of Subjects Who do not Meet the Case Definition (No Insomnia) and those Who do Meet the Case Definition (Insomnia)

	PSQI Global Score		Total	PSQI Global Score		Total
	≤5	>5		≤10	>10	
ISQ						
No Insomnia	93	137	230	207	23	230
Insomnia	2	24	26	15	11	26
Total	95	161	256	222	34	256

Additional composite variables were created to facilitate more similar comparisons between measures. The first new variable created a composite score from the sleep diaries which considered the presence or absence of a complaint and the frequency of the complaint (< 3 or ≥ 3 × per week) for one month. This composite score yielded a categorical score that more closely resembled the score from the ISQ. Since PSG only provides quantitative information for 3 days, a composite was not created from PSG. The ISQ does not differentiate individuals who have 1, 2, or 3 sleep symptoms. Moreover, the three variables used to identify a sleep symptom are highly collinear. So, we conducted a principal components analysis (PCA) on SOL, WASO, and SE (from diaries and PSG) to create a single factor devoid of units and overlap. A participant was identified as “positive” for insomnia if minimum quantitative criteria were met for SOL, WASO, or SE. Kappa statistics were used to test

agreement between findings from the proposed classification from the ISQ in relation to other sleep measures whose results had been dichotomized. These analyses were conducted with the use of SPSS Version 14.0.³⁶ A 2-sided probability of < 0.05 was designated as a statistically significant association.

RESULTS

IRT Analyses

Through a heuristic examination of ICC graphs, we identified a nocturnal sleep symptom frequency ≥ 3 times per week as the most discriminating cutpoint; this value also corresponds with established case definitions of insomnia^{10,24}(Figure 1a and 1b). For daytime symptoms, ICC graphs (not shown) indicated that being bothered “quite a bit” or “extremely” for any of the daytime consequences optimally distinguished this DSM-IV specific criterion among those believed to have insomnia from those without insomnia. Our selection of the cutpoint locations for the individual items was based on our visual inspection of the locations of the threshold parameters and the graphs of the information functions.

IRT analyses of the ISQ culminated in a scoring algorithm that identified insomnia based on the following responses: (1) presence of at least one of 3 sleep symptoms: difficulty initiating sleep, difficulty maintaining sleep, or unrefreshing sleep; (2) the symptom(s) occur with a minimum frequency of ≥ 3 times per week; (3) duration of the sleep symptom is ≥ 4 weeks; and (4) at least one aspect of daily life (e.g., difficulties at work

Table 2b—Cross-Tabulations Between SOL from Sleep Diary (≥ 31 minutes) and PSG (cutoff ≥ 31 minutes) and ISQ Classifications of Subjects Who do not Meet the Case Definition (No Insomnia) and Those Who do Meet the Case Definition (Insomnia)

	SOL (sleep diary)		Total	SOL (PSG)		Total
	< 31 min or ≥ 31 min but < 3x/week	≥ 31 min & ≥ 3 x/week#		< 31 min	≥ 31 min	
ISQ						
No Insomnia	236	1	237	203	37	240
Insomnia	24	2	26	20	6	26
Total	260	3	263	223	42	266

Data from sleep diaries were assessed for sleep complaint of ≥ 31 minutes and frequency of ≥ 3 x/week to identify parameters more closely resembling DSM-IV and RDC diagnostic criteria.

Table 2c—Cross-Tabulations Between WASO from Sleep Diary (≥ 31 minutes) and PSG (cutoff ≥ 31 minutes) and ISQ Classifications of Subjects Who do not Meet the Case Definition (No Insomnia) and Those Who do Meet the Case Definition (Insomnia)

	WASO (sleep diary)		Total	WASO (PSG)		Total
	< 31 min or ≥ 31 min but < 3x/week	≥ 31 min & ≥ 3 x/week#		< 31 min	≥ 31 min	
ISQ						
No Insomnia	232	3	235	55	185	240
Insomnia	23	3	26	5	21	26
Total	255	6	261	60	206	266

Data from sleep diaries were assessed for sleep complaint of ≥ 31 minutes and frequency of ≥ 3 x/week to identify parameters more closely resembling DSM-IV and RDC diagnostic criteria.

Table 2d—Cross-Tabulations Between Sleep Efficiency from Sleep Diary and PSG (cutoff of $< 85\%$) and ISQ Classifications of Subjects Who do not Meet the Case Definition (No Insomnia) and Those Who do Meet the Case Definition (Insomnia)

	Sleep Efficiency (sleep diary)		Total	Sleep Efficiency (PSG)		Total
	$\geq 85\%$ or < 85% but < 3x/week	< 85% & ≥ 3 x/week		$\geq 85\%$	< 85%	
ISQ						
No Insomnia	232	5	237	142	98	240
Insomnia	25	1	26	10	16	26
Total	257	6	263	152	114	266

Table 2e—Cross-Tabulations Between any Sleep Symptom and ISQ Classifications of Subjects Who do not Meet the Case Definition (No Insomnia) and Those Who do Meet the Case Definition (Insomnia)

	Any Sleep Symptom and Frequency ≥ 3 x/week (sleep diary)		Total	Any Sleep Symptom (PSG)		Total
	No	Yes		No	Yes	
ISQ						
No Insomnia	230	7	237	47	193	240
Insomnia	21	5	26	4	22	26
Total	251	12	263	51	215	266

or in social life) is affected “quite a bit” or “extremely” by endorsed sleep symptoms (Appendix 2). As a validity and sensitivity check, we considered additional frequency cutoffs with respect to each sleep complaint and daytime consequence to explore for variability in the sensitivity of the ISQ. Changing the cutpoint modestly affected, but did not improve, the sensitivity of the ISQ in the number of women identified as having insomnia (data not shown).

Sleep Characteristics

In the present sample, 38 (10.6%) women met diagnostic criteria for insomnia as defined by the ISQ. The subjective and PSG sleep measures for participants by ISQ insomnia criteria are reported in Table 1. Among participants without clinically significant sleep disordered breathing ($AHI < 15$), 26 (9.8%) of the cohort met ISQ-derived criteria for insomnia, indicat-

Table 3—Evaluation of the Insomnia Symptom Questionnaire as a Diagnostic Screening Tool

Concurrent Confirmation Measure	Sensitivity	Specificity	PPV	NPV	Kappa	LR
PSQI (> 5)	14.9%	97.9%	92.3%	40.4%	0.10	1.55
PSQI (> 10)	32.4%	93.2%	42.3%	90.0%	0.28	4.23
SOL – diary (≥ 31 minutes) & $>3x/week$	66.7%	90.8%	7.7%	99.6%	0.12	3.44
WASO – diary (≥ 31 minutes) & $>3x/week$	50.0%	91.0%	11.5%	98.7%	0.16	9.01
SE – diary ($< 85\%$) & $>3x/week$	16.7%	90.3%	3.8%	97.9%	0.03	1.82
Any sleep symptom & $> 3x/week$ (Diary)	41.7%	91.6%	19.2%	97.0%	0.21	6.52
SOL – PSG (≥ 31 minutes)	14.3%	91.0%	23.1%	84.6%	0.05	1.38
WASO – PSG (≥ 31 minutes)	10.2%	91.7%	80.8%	22.9%	0.01	1.04
SE – PSG ($< 85\%$)	14.0%	93.4%	61.5%	59.2%	0.08	1.51

PPV = Positive predictive value; NPV = Negative predictive value; LR = Likelihood ratio for a positive test; PSQI = Pittsburgh Sleep Quality Index; SOL = Sleep onset latency; WASO = Wake after sleep onset; SE = Sleep Efficiency; PSG = Polysomnography; Any sleep complaint = endorsement of SOL ≥ 31 min OR WASO ≥ 31 min OR SE $< 85\%$

ing symptom overlap between symptoms of insomnia and sleep apnea. Women who were identified as having insomnia on the ISQ had more sleep symptoms (SOL ≥ 31 minutes or WASO ≥ 31 minutes or SE $< 85\%$) indicating greater sleep disturbance, as measured by sleep diary ($F_{(261)} = 8.3$, $p = 0.001$) and PSG measures ($F_{(264)} = 12.2$, $p = 0.02$), than did women who were not identified as having insomnia on the ISQ.

Psychometric and Test Characteristics

Content and face validity were considered excellent since the ISQ questions were derived directly from widely accepted and published criteria from DSM-IV¹⁰ and RDC.²⁴ The 13-item ISQ had an overall reliability coefficient (Cronbach α) of 0.89, indicating a high degree of internal consistency.³²

The criterion validity of the ISQ was assessed by examining the diagnostic accuracy of ISQ outcomes referenced to dichotomized PSQI scores and dichotomized estimates of SOL, WASO, and SE, and subsequent composites, from sleep diary and PSG (Table 2a-e). Tables 2b and 2c indicate that comparing the ISQ to a composite measure derived from the sleep diary consisting of SOL or WASO ≥ 31 minutes or SE $< 85\%$ and frequency of ≥ 3 times per week results in more true negatives than quantitative data from PSG alone. The highest sensitivities (66.7% and 50.0%), and specificities (90.8% and 91.0%), with respect to SOL or WASO, were from the sleep diary (Table 3). When comparing the dichotomized ISQ to other sleep measures, specificities were quite high (all $> 90\%$), indicating that the ISQ identifies those without insomnia in a manner consistent with sleep parameters from the sleep diaries, PSG, and PSQI. Sensitivities (10.1% to 66.7%) were much more variable, indicating that the ISQ insomnia definition was not sensitive to the manner in which those sleep measures characterize insomnia. The wide range of PPV similarly indicates that traditional sleep measures and methodologies differentially identified false positives. Similarly, NPV was highest when composite measures were assessed (90% to 99%). With respect to the composite variables, the sleep diary composite variable (which includes any sleep symptom and minimum frequency ≥ 3 times per week) most closely resembled the ISQ definition as it considers 2 of the 4 criteria used to diagnose insomnia. PSG, on the other hand, can only capture one criterion, i.e., the presence of sleep disturbances, and it compared poorly with the ISQ.

As seen in Table 3, the kappa values between dichotomized ISQ scores and PSQI scores, as well as cutoffs derived from the

sleep diary and PSG variables (SOL ≥ 31 minutes, WASO ≥ 31 minutes, and SE $< 85\%$), indicate no agreement when compared to reference values proposed by Landis.³⁷ The LRs indicate that the ISQ is more similar to the composite measure from the sleep diary that includes WASO and frequency ≥ 3 times per week in its ability to identify who has insomnia (LR > 6). The single diary SE measure and all PSG data do not provide useful measures of validity (LRs 1-2) (Table 3).

DISCUSSION

There is a growing need for an easily administered and scored self-report measure that provides a clinically relevant case-definition of insomnia guided by established diagnostic criteria, such as the DSM-IV¹⁰ and the AASM's RDC.²⁴ The ISQ was developed to provide such a tool. We used IRT analysis to evaluate which ISQ items provided the most information regarding symptoms of insomnia and which cutoff on those items would support the specification of a simple scoring algorithm with a dichotomous outcome of insomnia ("present" or "absent") that is suitable for use in large scale research studies. Our preliminary evaluation suggests that the ISQ effectively identifies insomnia caseness based on DSM-IV¹⁰ and RDC²⁴ criteria. In contrast to other self-report measures, the ISQ contributes additional information for complaint criteria of ≥ 3 times per week and at least one month duration, respectively, in the diagnosis process.

The case definition of insomnia includes an aggregate of symptoms together with severity.^{10,24} In epidemiological or health services research, and in research settings where trained sleep medicine specialists are not available, the use of a structured clinical interview is impractical and costly.³⁸ Since few current retrospective instruments assess all recommended criteria,^{10,24} particularly frequency or daytime consequences, the use of multiple complementary, reliable methods, including retrospective and prospective estimates, is typically suggested to adequately capture all the different dimensions.^{11,38} In the context of a research study, when clinical interviews are not an option, a combination of other proxies, such as patient history, physical exam, sleep diaries, and questionnaires, can generate a reliable estimate of insomnia caseness.^{38,33,11,39} That the ISQ captures a defined set of insomnia symptoms enhances its interpretability in research studies when a clinical interview is not feasible.

Insomnia is a multidimensional disorder.³⁸⁻⁴¹ Since the ISQ assesses the recommended criteria, while traditional sleep mea-

asures assess one or two criteria, it is not necessarily surprising that there was low agreement for the diagnosis of insomnia between the ISQ and the other measures of sleep using kappa statistics.⁴² Likewise, the low LRs and variable PPV suggest that if a symptom cluster more closely approximates the DSM-IV¹⁰ or RDC²⁴ diagnosis, then the gold standard should include symptom frequency criteria. The ISQ appears superior to current methods since it assesses the multidimensionality of the insomnia disorder better than traditional methods.

Sleep diaries are commonly used to provide a detailed, prospective account about quantitative sleep disturbances and their variability.³⁸ However, quantitative information alone is insufficient to diagnose insomnia. To allow a more reliable comparison between the sleep diary and the ISQ, we utilized the recommended cutoff of ≥ 31 minutes in conjunction with an associated frequency. We are not aware of any other published reports that have measured frequency of a sleep complaint when using sleep diaries as comparison measures.¹⁶ This cutoff as compared with ≥ 30 minutes or ≥ 60 minutes appears to provide a more meaningful estimate and one that better reflects an unpleasant sleep experience.¹¹

Validating an insomnia questionnaire like the ISQ is best served in a population that experiences high rates of insomnia so that there are a sufficient number of true positives. Using the SWAN Sleep Study to validate the ISQ was ideal since many women going through the menopausal transition report symptoms of insomnia.^{6,7,43,44} The ISQ identified an insomnia diagnosis prevalence of 9.8% in this cohort of pre-, peri, and postmenopausal women, excluding those with an AHI ≥ 15 . This prevalence rate is consistent with the prevalence rates reported by Ohayon (2002) which were 6% for the general population, corresponding to a decision-making diagnosis using DSM-IV classifications of criteria for insomnia, and 9% to 15% when the diagnosis was based on insomnia symptoms and daytime consequences.⁴⁵ Our prevalence estimate is also similar to a 14.3% prevalence in a previous reported among menopausal Korean women who were asked about difficulty initiating or maintaining sleep or early morning awakenings.⁴⁶

A limitation of the current study is the absence of a clinical diagnosis as a gold standard for validation. Future studies are needed to compare the ISQ to clinical interviews, evaluate the ISQ's test-retest reliability, and assess other measures of validity. Consistent with information gleaned during diagnostic interviews, it might prove useful to include question(s) regarding opportunity/circumstances for sleep as the ISQ undergoes further refinement. Inclusion of this item may improve its sensitivity since it is a criterion that is specific for the diagnosis of insomnia. The ISQ does not assess insomnia due to another medical or mental condition and since comorbidity was not evaluated in this report, it is possible that some of the women may have insomnia due to another medical or mental condition. A follow-up evaluation of the ISQ might take into account comorbid conditions such as depression or anxiety. Lastly, sleep symptom reporting and prevalence rates for insomnia differ between men and women, thus a limitation is that this sample was all women.

A distinctive quality of the ISQ is that it is not only a severity measure for insomnia, but it is an instrument that appears to provide an accurate case definition of insomnia. Following further psychometric evaluations, we believe that the ISQ is likely to

emerge as a cost-effective self-report instrument with high utility for use in large observational studies as a screening instrument.

ACKNOWLEDGMENTS

The Study of Women's Health Across the Nation (SWAN) has grant support from the National Institutes of Health (NIH), DHHS, through the National Institute on Aging (NIA), the National Institute of Nursing Research (NINR) and the NIH Office of Research on Women's Health (ORWH) (Grants NR004061; AG012505, AG012535, AG012531, AG012539, AG012546, AG012553, AG012554, AG012495). The content of this article manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NIA, NINR, ORWH or the NIH.

Funding for the SWAN Sleep Study is from the National Institute on Aging (Grants AG019360, AG019361, AG019362, AG019363).

Clinical Centers: University of Michigan, Ann Arbor - Mary-Fran Sowers, PI; Massachusetts General Hospital, Boston, MA - Robert Neer, PI 1994 - 1999; Joel Finkelstein, PI 1999- present; Rush University, Rush University Medical Center, Chicago, IL - Lynda Powell, PI; University of California, Davis/Kaiser - Ellen Gold, PI; University of California, Los Angeles - Gail Greendale, PI; University of Medicine and Dentistry - New Jersey Medical School, Newark -Gerson Weiss, PI 1994 - 2004; Nanette Santoro, PI 2004 - present; and the University of Pittsburgh, Pittsburgh, PA - Karen Matthews, PI.

NIH Program Office: National Institute on Aging, Bethesda, MD - Marcia Ory 1994 - 2001; Sherry Sherman 1994 - present; National Institute of Nursing Research, Bethesda, MD - Program Officers.

Central Laboratory: University of Michigan, Ann Arbor - Daniel McConnell (Central Ligand Assay Satellite Services).

Coordinating Center: New England Research Institutes, Watertown, MA - Sonja McKinlay, PI 1995 - 2001; University of Pittsburgh, Pittsburgh, PA - Kim Sutton-Tyrrell, PI 2001 - present.

Steering Committee: Chris Gallagher, Chair, Susan Johnson, Chair

N-CTRC grant RR024153 for support of the conduct sleep studies including scoring, reliability assessments, and processing of sleep data. Dr. Moul's time is supported by grants MH 24652-29,1 U01 AR052155-01, and 5 P01 AG20677-02. Dr. Okun's time is supported by NR010813.

DISCLOSURE STATEMENT

This is not an industry supported study. Dr. Buysse has consulted for Actelion, Arena, Cephalon, Eli Lilly, GlaxoSmithKline, Merck, Neurocrine, Neurogen, Pfizer, Respiroics, Sanofi-Aventis, Sepracor, Servier, Stress Eraser, Takeda, and Transcept Pharmaceuticals, Inc. The other authors have indicated no financial conflicts of interest.

REFERENCES

1. National Institutes of Health State of the Science Conference statement on Manifestations and Management of Chronic Insomnia in Adults, June 13-15,2005. *Sleep* 2005;28:1049-57.
2. Krystal AD. Insomnia in women. *Clin Cornerstone* 2003;5:41-50.

3. Ohayon MM. Epidemiological study on insomnia in a general population. *Sleep* 1996;19:S7-S15.
4. Leger D, Guilleminault C, Dreyfus JP, Delahaye C, Paillard M. Prevalence of insomnia in a survey of 12,778 adults in France. *J Sleep Res* 2000;9:35-42.
5. Zhang B, Wing YK. Sex differences in insomnia: a meta-analysis. *Sleep* 2006;29:85-93.
6. Kravitz HM, Ganz PA, Bromberger J, Powell LH, Sutton-Tyrrell K, Meyer PM. Sleep difficulty in women at midlife: a community survey of sleep and the menopausal transition. *Menopause* 2003;10:19-28.
7. Krystal AD, Edinger J, Wohlgemuth W, Marsh GR. Sleep in peri-menopausal and post-menopausal women. *Sleep Med Rev* 1998;2:243-53.
8. Owens JF, Matthews KA. Sleep disturbance in healthy middle-aged women. *Maturitas* 1998;30:41-50.
9. American Academy of Sleep Medicine. The International Classification of Sleep Disorders, Second Edition (ICSD-2): Diagnostic and Coding Manual. Second Edition ed. 2005.
10. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-IV-TR). 4th ed. Text Revision ed. Washington, DC: American Psychiatric Association; 2000.
11. Lichstein KL, Durrence HH, Taylor DJ, Bush AJ, Riedel BW. Quantitative criteria for insomnia. *Behav Res Ther* 2003;41:427-45.
12. Schramm E, Hohagen F, Grasshoff U, et al. Test-retest reliability and validity of the Structured Interview for Sleep Disorders According to DSM-III-R. *Am J Psychiatry* 1993;150:867-72.
13. Edinger JD. Classifying insomnia in a clinically useful way. *J Clin Psychiatry* 2004;65 Suppl 8:36-43.
14. Reite M, Buysse D, Reynolds C, Mendelson W. The use of polysomnography in the evaluation of insomnia. *Sleep* 1995;18:58-70.
15. Moul DE, Hall M, Pilkonis PA, Buysse DJ. Self-report measures of insomnia in adults: rationales, choices, and needs. *Sleep Med Rev* 2004;8:177-98.
16. Bastien CH, Vallieres A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med* 2001;2:297-307.
17. Hays RD, Martin SA, Sesti AM, Spritzer KL. Psychometric properties of the Medical Outcomes Study Sleep measure. *Sleep Med* 2005;6:41-4.
18. 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.
19. 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.
20. Violani C, Devoto A, Lucidi F, Lombardo C, Russo PM. Validity of a short insomnia questionnaire: the SDQ. *Brain Res Bull* 2004;63:415-21.
21. 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.
22. Backhaus J, Jughanns 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):737-40.
23. Sowers MF, Crawford SL, Sternfeld B, et al. SWAN: A multi-center, multiethnic, community-based cohort study of women and the menopausal transition. *Menopause: biology and pathobiology*. Academic Press; 2000. p. 175-88.
24. Edinger JD, Bonnet MH, Bootzin RR, et al. Derivation of research diagnostic criteria for insomnia: report of an American Academy of Sleep Medicine Work Group. *Sleep* 2004;27:1567-96.
25. Monk TH, Reynolds CF, Kupfer DJ, et al. The Pittsburgh Sleep Diary. *J Sleep Res* 1994;3:111-20.
26. Rechtschaffen A, Kales A. A Manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. NIH Publication 204. Washington, DC: U.S. Government Printing Office, Department of Health Education and Welfare; 1968.
27. American Academy of Sleep Medicine Task Force. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. *Sleep* 1999;22:667-89.
28. American Sleep Disorders Association Atlas Task Force. Recording and scoring leg movements. *Sleep* 1993;16:749-59.
29. Baker F. The Basics of Item Response Theory. College Park: ERIC Clearinghouse on Assessment and Evaluation; 2001.
30. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. *Med Care* 2000;38:II28-II42.
31. MULTILOG user's guide: multiple categorical item analysis and test scoring using item response theory. [computer program]. Version 7.0 Scientific Software International; 2003.
32. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-309.
33. Faraone SV, Tsuang MT. Measuring diagnostic accuracy in the absence of a "gold standard". *Am J Psychiatry* 1994;151:650-7.
34. Flemons WW, Whitelaw WA, Brant R, Remmers JE. Likelihood ratios for a sleep apnea clinical prediction rule. *Am J Respir Crit Care Med* 1994;150:1279-85.
35. Lineberger MD, Carney CE, Edinger JD, Means MK. Defining insomnia: quantitative criteria for insomnia severity and frequency. *Sleep* 2006;29:479-85.
36. SPSS 14.0 for Windows [computer program]. Chicago: SPSS; 2006.
37. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977;33:363-74.
38. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. *Sleep* 2006;29:1155-73.
39. Summers MO, Crisostomo MI, Stepanski EJ. Recent developments in the classification, evaluation, and treatment of insomnia. *Chest* 2006;130:276-86.
40. Chesson A Jr, Hartse K, Anderson WM, et al. Practice parameters for the evaluation of chronic insomnia. An American Academy of Sleep Medicine report. Standards of Practice Committee of the American Academy of Sleep Medicine. *Sleep* 2000;23:237-41.
41. Sateia MJ, Doghramji K, Hauri PJ, Morin CM. Evaluation of chronic insomnia. An American Academy of Sleep Medicine review. *Sleep* 2000 15;23:243-308.
42. Gordis L. Assessing the validity and reliability of diagnostic and screening tests. *Epidemiology*. 2nd ed. Philadelphia: W.B. Saunders; 2000:63-81.
43. Chung KF, Tang MK. Subjective sleep disturbance and its correlates in middle-aged Hong Kong Chinese women. *Maturitas* 2006 20;53:396-404.
44. Young T, Rabago D, Zgierska A, Austin D, Laurel F. Objective and subjective sleep quality in premenopausal, perimenopausal, and postmenopausal women in the Wisconsin Sleep Cohort Study. *Sleep* 2003;26:667-72.
45. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. *Sleep Med Rev* 2002;6:97-111.
46. Shin C, Lee S, Lee T, et al. Prevalence of insomnia and its relationship to menopausal status in middle-aged Korean women. *Psychiatry Clin Neurosci* 2005;59:395-402.

Appendix 1—Insomnia Symptom Questionnaire (ISQ)

Name _____ ID# _____ Date _____ Age _____

Instructions: If you have experienced **any** sleep symptoms **during the past month** please circle the appropriate number to let us know how your sleep is affecting your daily life.

During the past month did you have...	Never	Do not know	Rarely less than once per week	Sometimes 1-2 times per week	Frequently 3-4 times per week	Always 5-7 times per week	How long has the symptom lasted? (# of weeks, months or years)
1. Difficulty falling asleep?	0	1	2	3	4	5	<input type="checkbox"/> wks <input type="checkbox"/> mos ____ <input type="checkbox"/> yrs
2. Difficulty staying asleep?	0	1	2	3	4	5	<input type="checkbox"/> wks <input type="checkbox"/> mos ____ <input type="checkbox"/> yrs
3. Frequent awakenings from sleep?	0	1	2	3	4	5	<input type="checkbox"/> wks <input type="checkbox"/> mos ____ <input type="checkbox"/> yrs
4. Feeling that your sleep is not sound?	0	1	2	3	4	5	<input type="checkbox"/> wks <input type="checkbox"/> mos ____ <input type="checkbox"/> yrs
5. Feeling that your sleep is unrefreshing?	0	1	2	3	4	5	<input type="checkbox"/> wks <input type="checkbox"/> mos ____ <input type="checkbox"/> yrs

Instructions: If you have experienced **any** sleep symptoms **during the past month** please circle the appropriate number to let us know how your sleep is affecting your daily life

During the past month...	Not at all	A little bit	Moderately	Quite a bit	Extremely
6. How much do your sleep problems bother you?	0	1	2	3	4
7. Have your sleep difficulties affected your work?	0	1	2	3	4
8. Have your sleep difficulties affected your social life?	0	1	2	3	4
9. Have your sleep difficulties affected other important parts of your life?	0	1	2	3	4
10. Have your sleep difficulties made you feel irritable?	0	1	2	3	4
11. Have your sleep problems caused you to have trouble concentrating?	0	1	2	3	4
12. Have your sleep difficulties made you feel fatigued?	0	1	2	3	4
13. How sleepy do you feel during the day?	0	1	2	3	4

Appendix 2—ISQ Scoring Algorithm

The Insomnia Symptom Questionnaire (ISQ) has 13 self-rated questions. Only questions 1,2 or 5 are used to determine the presence, frequency AND duration of sleep symptom criteria. Questions 6-13 are used to identify significant daytime consequences of the sleep complaint. Please answer the following questions based on the participants responses to determine if they meet the case definition of insomnia. Scoring proceeds as follows:

	Yes	No
Sleep symptom criterion		
Is the answer to at least one of Questions 1, 2, or 5 “Frequently” or “Always”?	<input type="checkbox"/>	<input type="checkbox"/>
Duration criterion		
Is the answer to at least one of Questions 1, 2, or 5 “ \geq 4 weeks”?	<input type="checkbox"/>	<input type="checkbox"/>
Daytime impairment criterion		
Is the answer to at least one of Questions 6 -13 “Quite a bit” or “Always”?	<input type="checkbox"/>	<input type="checkbox"/>
If the answer to each question above is “Yes,” assign a diagnosis of insomnia disorder.	Insomnia disorder <input type="checkbox"/>	
If the answer to one or more of the questions above is “No,” do not assign a diagnosis of insomnia disorder.	No insomnia disorder <input type="checkbox"/>	