The Insomnia Severity Index: Psychometric Indicators to Detect Insomnia Cases and Evaluate Treatment Response

Charles M. Morin, PhD; Geneviève Belleville, PhD; Lynda Bélanger, PhD; Hans Ivers, PhD

Université Laval and Centre de recherche Université Laval-Robert Giffard, Québec, Canada

Background: Although insomnia is a prevalent complaint with significant morbidity, it often remains unrecognized and untreated. Brief and valid instruments are needed both for screening and outcome assessment. This study examined psychometric indices of the *Insomnia Severity Index* (ISI) to detect cases of insomnia in a population-based sample and to evaluate treatment response in a clinical sample.

Methods: Participants were 959 individuals selected from the community for an epidemiological study of insomnia (Community sample) and 183 individuals evaluated for insomnia treatment and 62 controls without insomnia (Clinical sample). They completed the ISI and several measures of sleep quality, fatigue, psychological symptoms, and quality of life; those in the Clinical sample also completed sleep diaries, polysomnography, and interviews to validate their insomnia/good sleep status and assess treatment response. In addition to standard psychometric indices of reliability and validity, item response theory analyses were computed to examine ISI item response patterns. Receiver operating curves were used to derive optimal cutoff scores for case identification and to quantify the minimally important changes in relation to global improvement ratings obtained by an independent assessor.

Results: ISI internal consistency was excellent for both samples (Cronbach α of 0.90 and 0.91). Item response analyses revealed adequate discriminatory capacity for 5 of the 7 items. Convergent validity was supported by significant correlations between total ISI score and measures of fatigue, quality of life, anxiety, and depression. A cutoff score of 10 was optimal (86.1% sensitivity and 87.7% specificity) for detecting insomnia cases in the community sample. In the clinical sample, a change score of −8.4 points (95% CI: −7.1, −9.4) was associated with moderate improvement as rated by an independent assessor after treatment.

Conclusion: These findings provide further evidence that the ISI is a reliable and valid instrument to detect cases of insomnia in the population and is sensitive to treatment response in clinical patients.

Keywords: Insomnia, assessment, evaluation, measure, patient-reported outcome

Citation: Morin CM; Belleville G; Bélanger L; Ivers H. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. *SLEEP* 2011;34(5):601-608.

INTRODUCTION

Insomnia is a highly prevalent condition and carries significant burden in terms of functional impairment, health care costs, and increased risk of depression.[1-7] Despite its high prevalence and significant morbidity, insomnia often remains unrecognized and untreated, partly due to several barriers to assessment. Accurate case identification is important for deriving valid estimates of prevalence/incidence and for assessing burden of disease in the population. Identifying clinically significant insomnia is also important to intervene early and reduce morbidity. Thus, reliable and valid instruments are needed to assist investigators and clinicians in evaluating insomnia in various research and clinical contexts.

The assessment of insomnia is multidimensional and should ideally include a clinical evaluation and be complemented by self-report questionnaires and daily sleep diaries. While a clinical evaluation remains the gold standard for making a valid insomnia diagnosis,[8,9] such an evaluation can be time-consuming in routine clinical practice and may discourage some health practitioners from systematically inquiring about sleep in all of their patients. Brief and valid questionnaires can facilitate

the initial screening and formal evaluation of insomnia. The patient's perspective is also of critical importance to monitor progress and evaluate outcome after initiating treatment. From a regulatory perspective, patient-reported outcomes are becoming increasingly used to substantiate evidence of treatment effectiveness in clinical trials. There is a need for assessment tools that are brief, practical, and psychometrically sound both for screening purposes and treatment outcome evaluation.

There are currently several patient-reported questionnaires available for assessing insomnia symptoms, severity, correlates, and a variety of constructs presumed to contribute to the etiology of insomnia.[8,10] With regard to screening for insomnia and evaluating treatment outcome, there are fewer choices available. Some of the most widely used instruments for these purposes include, for example, the Insomnia Severity Index,[11] the Pittsburgh Sleep Quality Index,[12] the Insomnia Symptom Questionnaire,[13] and the Athens Insomnia Scale.[14] While the number of items, response format, and time frame varies across instruments, they are generally aimed at assessing the patient's perception and at quantifying subjective dimensions of insomnia. Each of these instruments has its own advantages and limitations (for reviews see Buysse et al., Martin et al., Morin, and Moul et al.).[10,15-17] The *Insomnia Severity Index* (ISI) is a brief instrument that was designed to assess the severity of both nighttime and daytime components of insomnia. It is available in several languages and is increasingly used as a metric of treatment response in clinical research. While its psychometric properties using classical test theory have been documented previously,[11,18-20] the present paper reports further validation using item response theory (IRT) analyses to examine response

patterns on individual ISI items and receiver-operating curves (ROC) to identify optimal cut points for case finding in a community sample and for assessing treatment response in a clinical sample.

METHODS

Participants

Data were derived from several studies conducted in our sleep research center at Laval University. Data for individuals with insomnia were gathered from 2 studies, including a treatment study evaluating the efficacy of cognitive-behavior therapy, singly and combined with medication, for persistent insomnia,[21] and an epidemiological study documenting the prevalence, incidence, and natural history of insomnia in a population-based sample.[22] Data from controls without insomnia complaints were derived from the same epidemiological study and from other cross-sectional studies comparing individuals with insomnia and control good sleepers on psychological, neuropsychological, and health-related variables. All subjects were thus solicited to participate in a sleep/insomnia research study, with the main difference that some participants (Clinical sample) responded to media advertisements to participate in clinical studies of insomnia, whereas those from the Community sample included individuals selected from a population-based sample who were ≥ 18 years and volunteered to participate in the study. The only individuals excluded from this cohort were those who reported having been diagnosed with a sleep disorder other than insomnia. Two samples were thus formed for the present paper: a Community sample including 959 individuals (with and without insomnia) selected from the adult population to participate in the epidemiological study,[22,23] and a Clinical sample, including 183 individuals with insomnia enrolled in a treatment study[21] and 62 healthy controls (without insomnia) enrolled in one of several cross-sectional studies.

Measures

Insomnia Severity Index (ISI)

The ISI is a 7-item self-report questionnaire assessing the nature, severity, and impact of insomnia.[11,24] The usual recall period is the "last month" and the dimensions evaluated are: severity of sleep onset, sleep maintenance, and early morning awakening problems, sleep dissatisfaction, interference of sleep difficulties with daytime functioning, noticeability of sleep problems by others, and distress caused by the sleep difficulties. A 5-point Likert scale is used to rate each item (e.g., 0 = no problem; 4 = very severe problem), yielding a total score ranging from 0 to 28. The total score is interpreted as follows: absence of insomnia (0-7); sub-threshold insomnia (8-14); moderate insomnia (15-21); and severe insomnia (22-28). Three versions are available—patient, clinician, and significant others—but the present paper focuses on the patient version only. Previous studies have reported adequate psychometric properties for both the English and French versions.[11,20]

Pittsburgh Sleep Quality Index (PSQI)

The PSQI is a 19-item questionnaire evaluating sleep quality and disturbances over the past month.[12] The first 4 items are open questions, whereas items 5 to 19 are rated on a 4-point Likert scale. Individual items scores yield 7 components. A total score, ranging from 0 to 21, is obtained by adding the 7 component scores. A score > 5 suggests poor sleep quality. The psychometric properties of the PSQI have been documented in multiple studies,[25] including one with a French-Canadian sample.[20] The PSQI was used because it measures a construct (sleep quality) that is related to insomnia but a construct that is broader than insomnia severity. It was administered to the community sample only.

Multidimensional Fatigue Inventory (MFI)

The MFI is composed of 20 statements for which the responder has to indicate, on a 5-point Likert scale, to what extent the particular item applies to his or her situation in the previous days.[26] The MFI measures 5 dimensions of fatigue: general fatigue, mental fatigue, physical fatigue, reduced activity, and reduced motivation. For each scale, the score varies between 4 and 20, with a higher score indicating more severe fatigue. The internal consistency and construct validity of this scale are adequate; the general fatigue scale is the most sensitive subscale. This measure was included because fatigue is one of the most common daytime symptoms among individuals with insomnia and should be significantly related to insomnia severity.

Beck Depression Inventory (BDI)

The BDI is a 21-item questionnaire assessing (on 4-point Likert scales) the intensity of depressive symptoms in the past week.[27] The total score ranges from 0 to 63, with a higher score suggesting more severe depressive symptoms. This measure of depressive symptoms and those of anxiety symptoms were included because of the frequent co-occurrence of these psychological symptoms with insomnia.

Beck Anxiety Inventory (BAI)

The BAI is a 21-item questionnaire assessing, on 4-point Likert scales, the intensity of anxiety symptoms in the past week.[28] The total score ranges from 0 to 63, with higher scores indicating more severe anxiety symptoms. This questionnaire was completed by the Clinical sample only.

State-Trait Anxiety Inventory (STAI)

The STAI is a 2-part instrument assessing state (situational) and trait anxiety (in general).[29] Only the *Trait* part (STAI-Trait) was used in the present study. The STAI-Trait comprises 20 items rated on a 4-point Likert scale. Psychometric properties of the STAI are excellent. This instrument was completed by the Community sample only.

SF-12 Health Survey - version 2 (SF-12v2)

This instrument[30] is a short form of the SF-36, a widely used health survey. The 12 items are rated on a 5-point Likert scale, and 8 subscale scores can be derived: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health. Three subscales were used in the present study: general health, mental health, and bodily pain. The psychometric properties of the SF-12v2 are adequate with reliability coefficients for the 8 subscales ranging

from 0.73 to 0.87 in the general population.[31] This questionnaire was completed by the Community sample only.

Sleep diaries

Participants (with and without insomnia) in the Clinical sample completed a daily sleep diary[24] for a 2-week period as part of their initial evaluation. The following dependent variables were derived: sleep onset latency, wake after sleep onset, early morning awakening, total wake time, total sleep time, and sleep efficiency for examination of convergent validity with severity of specific insomnia symptoms rated on the ISI.

Polysomnography (PSG)

Participants (with and without insomnia) in the Clinical sample completed 3 nights of PSG recordings as part of their initial evaluation. Pooled data from the second and third nights were used to derive the main variables of total wake time, total sleep time, and sleep efficiency. The PSG montage included standard electroencephalographic (EEG), electromyographic (EMG), and electrooculographic (EOG) monitoring. Sleep stages were scored according to standard criteria.[32] Respiration (airflow, tidal volume, and oxygen saturation) and anterior tibialis EMG were recorded during the first night to rule out sleep apnea and periodic limb movement disorders.

Procedures

Participants from the Community sample were sent the sleep survey and other questionnaires by mail and asked to return them in a stamped, self-return envelope. Participants from the Clinical sample completed an initial telephone screening for eligibility and were then invited for a clinical evaluation during which they underwent a semi-structured interview to determine the presence or absence of insomnia[24] and the Structured Clinical Interview for DSM-IV (SCID)[33] to assess the presence of psychiatric disorders. They also completed the above-mentioned assessment. Participants enrolled in treatment were also evaluated at the end of therapy by an independent assessor, blind to treatment conditions, who provided a clinical global improvement rating of the degree of change since initiating treatment (0 = unchanged or worse; 1 = minimal improvement, did not alter the status of care needed for the patient; 2 = moderate improvement, partial remission of symptoms; 3 = marked improvement, complete or nearly complete remission of all symptoms).

Data Analyses

Data were checked for normality, outliers, and missing data. No imputation of missing data was performed. Data were analyzed using SPSS version 11, with a 2-tailed α level of 5%. Data for the Community and Clinical samples were analyzed separately. The ISI reliability was evaluated using the standard Cronbach α coefficient and item-total correlations for internal consistency. IRT analyses were computed on the 183 individuals with insomnia from the Clinical sample to examine response patterns on the different ISI items. Criterion validity was examined by comparing agreements between the classification of patients as insomniacs or good sleepers based on the original ISI cutoff scores (i.e., < 8, no insomnia, 8-14, sub-threshold insomnia, > 14, moderate to severe insomnia) and the classification of patients derived from an external criterion (i.e., the semi-structured interview for the Clinical sample and the response to a yes-no question from the survey assessing the presence of sleep difficulties for the Community sample). Convergent validity was assessed by computing Pearson correlation coefficients between ISI scores and sleep parameters derived from sleep diaries and PSG, and questionnaires measuring related, albeit different, constructs. Additional analyses using *receiver operating curves* (ROCs) were completed to estimate the sensitivity and specificity of a minimally important difference (MID) to detect clinical change after treatment.

RESULTS

Demographics

The Community sample comprised 959 individuals (60.1% women, mean age 43.8 y [SD 14.1], 24.7% with insomnia complaints). The Clinical sample included 183 individuals (61.2% women; mean age 50.7 y [SD 10.3]) with an insomnia diagnosis and 62 controls without insomnia (51.6% women; mean age 41.2 y [SD 18.7]). Participants from the Clinical sample were significantly older than those in the Community sample (48.4 vs. 43.9 y; t_{1196} = 4.48, P < 0.001), but there was no significant difference for gender distribution (39.8% and 41.2% of men in the 2 samples). Within the Clinical sample, there was also an age difference, with controls being younger than participants with insomnia, t_{243} = 3.81, P < 0.001. This difference was due to the age criterion (≥ 30 years old) used in the treatment study, while no such restriction was applied for the controls.

Reliability

High internal consistency coefficients were obtained for both the Community (Cronbach α = 0.90) and Clinical samples (Cronbach α = 0.91). Correlations between individual items and total ISI scores ranged from 0.55 to 0.81 (mean 0.71) for the Community sample and from 0.50 to 0.85 (mean 0.73) for the Clinical sample, suggesting that all items contributed significantly to the total score. Items about satisfaction and worry about sleep difficulties showed the highest item-total correlations (0.79-0.85), while those targeting insomnia symptoms such as difficulty falling asleep or early morning awakening showed lower correlations (0.50-0.66).

IRT Analyses

IRT analyses were computed to examine response patterns on the individual ISI items. IRT is based on the assumption that the probability to observe a response (no problem vs. mild vs. moderate vs. severe vs. very severe problem) on an item follows a mathematical function related to the person's endorsement of the construct of interest (i.e., insomnia severity).[34] IRT models are also known as latent trait models, suggesting that item responses are manifestations of the hypothesized construct which is not directly measured but mainly inferred. In polytomous IRT response models, where item responses are measured on an ordinal Likert scale, 2 item parameters are estimated including the slope of the mathematical function, which represent the discriminative capacity of the item, and the step size parameters (k -1 parameters for an item with k response choices),
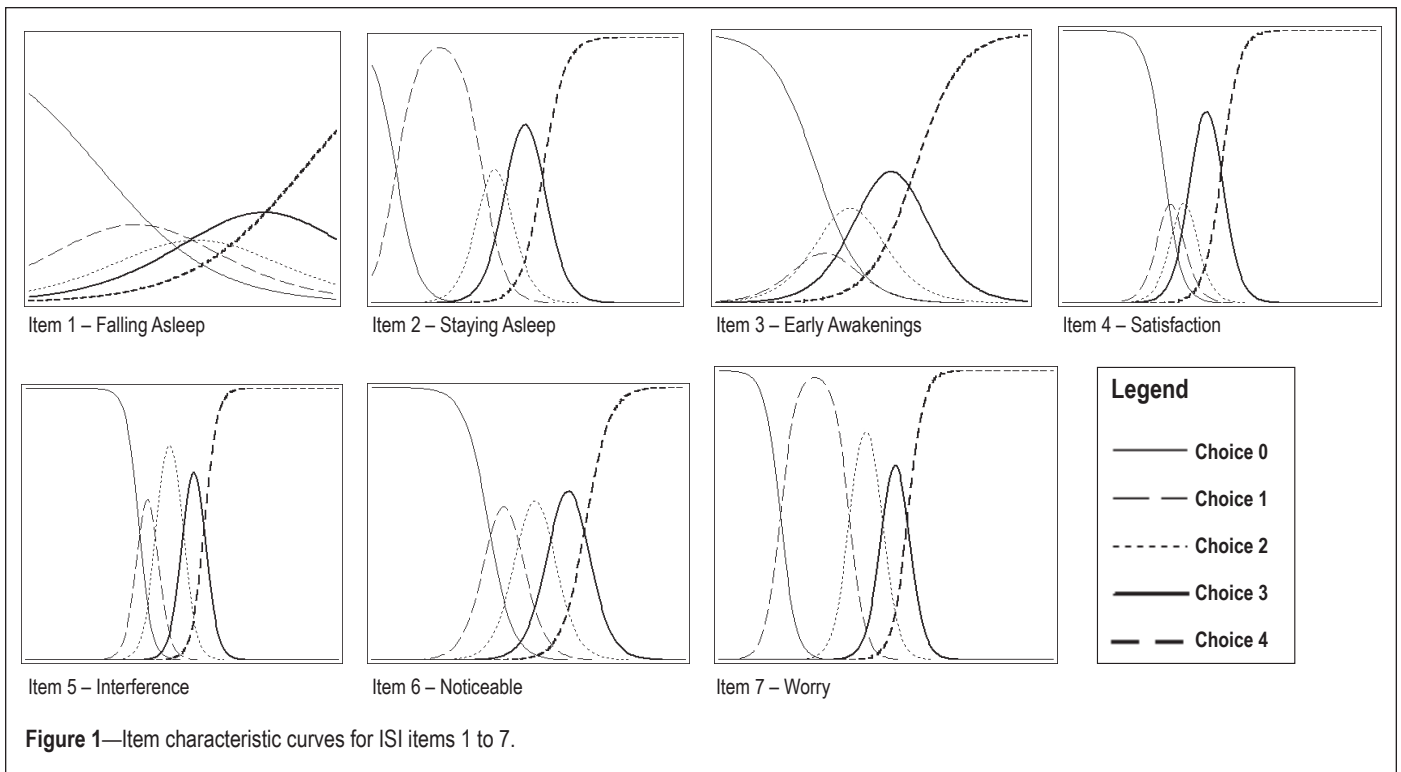
Figure 1—Item characteristic curves for ISI items 1 to 7.

Table 1—Percentage of clinical sample who endorsed each item response

Item ISI	Item response choice*				
	0	1	2	3	4
1. falling asleep	19.6	25.3	22.8	21.5	10.8
2. staying asleep	0.0	4.4	17.7	53.8	24.1
3. early awakening	9.5	8.2	27.9	39.9	14.6
4. satisfaction	0.6	1.9	6.3	43.7	47.5
5. interference	1.3	7.6	43.0	36.1	12.0
6. noticeable	6.3	29.1	45.6	17.1	1.9
7. worry	0.0	4.4	38.6	41.1	15.8

*Items 1-3 0, no problem; 1, mild; 2, moderate; 3, severe; 4, very severe
*Item 4   0, very satisfied; 1, satisfied; 2, neutral; 3, dissatisfied; 4, very dissatisfied
*Items 5-7 0, not at all; 1, a little; 2, somewhat; 3, much; 4, very much

which capture the location of the function curve for each item choice along the trait scale.

IRT analyses were completed on the 7 ISI items for the Clinical sample ($N = 183$ insomniacs) according to the graded response model.[35] Estimates for slope and step size parameters, as well as item characteristics curves, were computed using SAS NLMIXED procedure.[36] Results of the item characteristic curves are displayed in Figure 1. Statistical tests on IRT model parameters revealed a nonsignificant slope for item 1 (P = 0.12), as well as nonsignificant steps for item 1 (all steps had P values ranging from 0.11 to 0.12), item 2 (1st step, P = 0.65), and item 7 (1st step, P = 0.72). All other parameters were significant.

Table 1 shows the percentage of individuals endorsing each item. Choices 0 (no symptom) and 1 (mild severity/impact) were 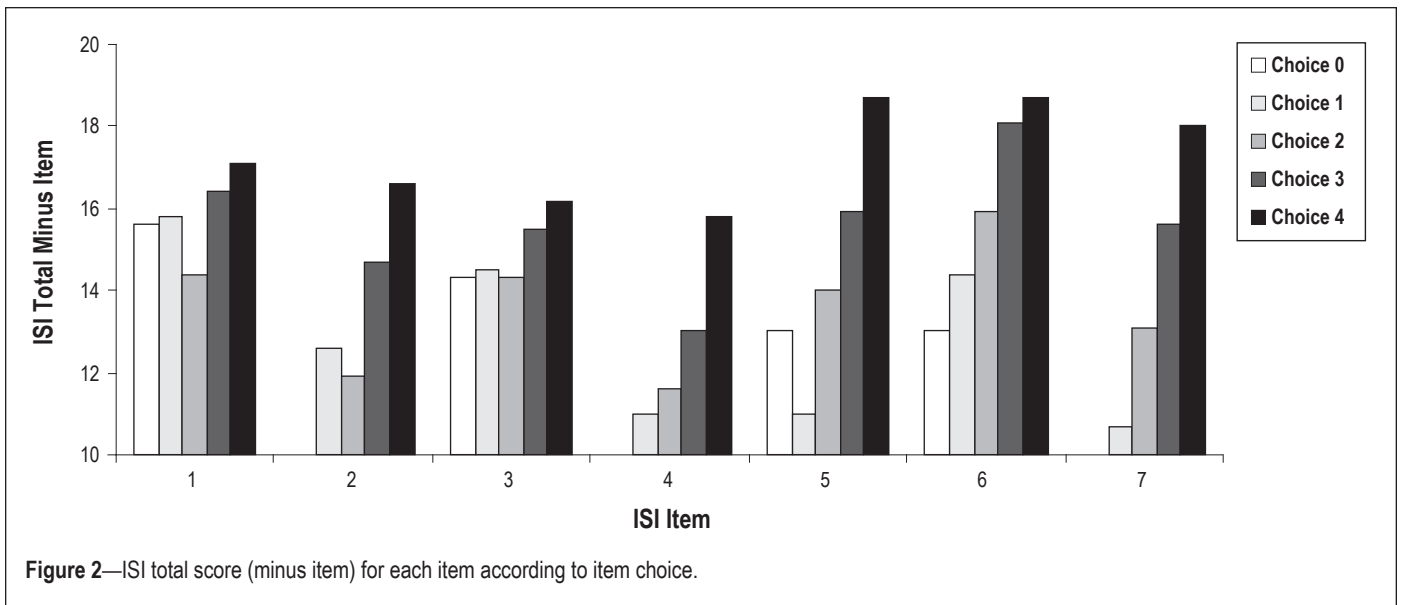used infrequently for items 2 (staying asleep), 5 (interference), and 7 (worry), while choices 1 to 3 were used infrequently for item 4 (sleep satisfaction). The impact of non-endorsement of an item is illustrated on the item characteristic curves (Figure 1); for example, for item 2 and 7, by the large bell curve observed for choice 1, indicating that choice 1 is the discriminative point for the lower dimension of the scale on these items. These results are not unexpected given the clinical nature of the sample (all participants met criteria for an insomnia diagnosis).

Figure 2 displays the averaged total ISI score according to each item response choice, after excluding the item score from the total score. Items 2, 4, 5, 6, and 7 show adequate to excellent discriminative capacity, as suggested by an almost linear increase of the total ISI score, as participants endorsed higher choices on these items. Item 1 shows poor discriminative capacity, as response choices are almost unrelated to total ISI score. This is also supported by the nonsignificant slope parameter for this item and the flat, large variability profile of its item characteristic curves (see Figure 1). Item 3 shows a somewhat similar response pattern but mostly for the lowest choices (0 to 2), a pattern that also translated into lower discriminative capacities (low and large bell curves) for these choices on the item characteristic curves.

Validity

Criterion validity

The ISI ability to identify the presence of insomnia was examined by studying the agreement between the proportion of patients classified as having insomnia based on the original ISI cutoff scores (i.e., < 8, no insomnia; > 14, moderate to severe insomnia) and the presence or absence of insomnia as determined by an external criterion. For the Clinical sample, this criterion was the diagnosis derived from the semi-structured

Figure 2—ISI total score (minus item) for each item according to item choice.

interview. As no interview was conducted with the Community sample, the criterion was the answer to a single yes/no question "*Do you consider that you have a sleep problem currently (last month)?*"

These analyses revealed that 98.3% of the Community sample reporting no sleep problem on the survey obtained ISI scores < 8 (no insomnia). Conversely, nearly 90% (89.7%) of individuals reporting a sleep problem on the survey obtained ISI scores ≥ 15 (moderate to severe insomnia). Individuals with scores at the subthreshold level (8-14) were about equally divided between those with (44%) and without sleep difficulties (56%). For the Clinical sample, 98.2% of participants without an insomnia diagnosis obtained ISI scores < 8, and none had an ISI score > 14. Conversely, 88.4% of those with an insomnia diagnosis obtained ISI scores ≥ 15.

Indices of sensitivity and specificity were derived for 3 different ISI cutoff scores (Table 2). A cutoff score of 8 (suggesting sub-threshold insomnia) was associated with sensitivity indices of 95.8% and 99.4% in the Community and Clinical samples, and with specificity indices of 78.3% and 91.8%, respectively. With a cutoff score of 15 (suggesting moderate to severe insomnia), specificity increased to 98.3% and 100% in the Community and Clinical samples, respectively, while sensitivity decreased to 47.7% and 78.1% for the 2 samples. A third cut-off score was derived empirically using ROC curves (not illustrated here) based on the actual results from the current

Table 2—Sensitivity and specificity data for the Clinical and Community samples

	Clinical Sample Cut-Off Score		
	ISI ≥ 8	ISI ≥ 11	ISI ≥ 15
Sensitivity	99.4% (177/178)	97.2% (173/178)	78.1% (139/178)
Specificity	91.8% (56/61)	100% (61/61)	100% (61/61)
False positive rate	8.2% (5/61)	0% (0/61)	0% (0/61)
False negative rate	0.6% (1/178)	2.8% (5/178)	21.9% (39/178)
Correct classification rate	97.5% (233/239)	97.9% (234/239)	83.7% (200/239)
Positive predictive power	97.3% (177/182)	100% (173/173)	100% (139/139)
Negative predictive power	98.2% (56/57)	92.4% (61/66)	61% (61/100)
	Community Sample Cut-Off Score		
	ISI ≥ 8	ISI ≥ 10	ISI ≥ 15
Sensitivity	95.8% (227/237)	86.1% (204/237)	47.7% (113/237)
Specificity	78.3% (565/722)	87.7% (633/722)	98.3% (710/722)
False positive rate	21.7% (157/722)	12.3% (89/722)	1.7% (12/722)
False negative rate	4.2% (10/237)	13.9% (33/237)	52.3% (124/237)
Correct classification rate	82.6% (792/959)	87.3% (657/959)	85.8% (823/959)
Positive predictive power	59.1% (227/384)	69.6% (204/293)	90.4% (113/125)
Negative predictive power	98.3% (565/575)	95.0% (633/666)	85.1% (710/834)

samples. For the Community sample, the highest correct classification rate was obtained with a cut point of 10, which led to 86.1% sensitivity and 87.7% specificity indices, respectively. For the Clinical sample, the highest correct classification rate was obtained with a cut point of 11, which yielded a 97.2% sensitivity and a perfect 100% specificity.

Convergent validity

Correlation coefficients were computed between scores on the first 3 ISI items (severity of initial, middle, and late insomnia) and the corresponding variables derived from daily sleep diaries or PSG (see Table 3). These data, obtained from the Clinical sample only, showed that the perceived severity of each insomnia symptom on the ISI was positively correlated with the corresponding diary variable but not significantly with most PSG variables. Total ISI score was also positively correlated with total wake

Table 3—Correlations between ISI and sleep parameters (from Clinical sample)

	Insomnia Severity Items			
	Difficulty Falling Asleep (Item #1)	Difficulty Staying Asleep (Item #2)	Waking Up Too Early (Item #3)	Total ISI Score
Sleep Diary				
Sleep onset latency	0.56*			
Wake after sleep onset		0.57*		
Number of awakenings		0.14*		
Early morning awakening			0.32*	
Total wake time				0.59*
Total sleep time				−0.54*
Sleep efficiency				−0.59*
Polysomnography				
Sleep onset latency	0.11			
Wake after sleep onset		0.14		
Number of awakenings		0.17*		
Early morning awakening			0.27	
Total wake time				0.11
Total sleep time				−0.13
Sleep efficiency				−0.16*

*P < 0.05

Table 4—Correlations between ISI scores and other questionnaires measuring other constructs

	ISI Total Score	
	Clinical sample	Community sample
Depression (BDI)	0.48*	0.50*
Anxiety (BAI)	0.48*	N/A
Anxiety (STAI-Trait)	N/A	0.50*
General fatigue (MFI)	0.41*	0.55*
Physical fatigue (MFI)	0.27*	0.45*
Reduced activities (MFI)	0.23*	0.31*
Reduced motivation (MFI)	0.20*	0.38*
Mental fatigue (MFI)	0.24*	0.37*
SF-12 Physical Health Component	N/A	−0.30*
SF-12 Mental Health Component	N/A	−0.51*
Sleep Quality (PSQI)	N/A	0.80*

BDI, Beck Depression Inventory; BAI, Beck Anxiety Inventory; MFI, Multidimensional Fatigue Inventory; SF-12, 12-Item Short-Form Health Survey; PSQI, Pittsburgh Sleep Quality Index. *P < 0.05.

Table 5—ISI change scores (baseline to post-treatment) according to clinical global improvement assessment (Clinical sample)

Clinical global improvement (external evaluator)	N	Mean (change scores)	95% CI
Slight improvement	20	−4.65	−2.61 to −6.69
Moderate improvement	63	−8.36	−7.20 to −9.53
Marked improvement	63	−9.89	−8.74 to −11.04

time and negatively correlated with total sleep time and sleep efficiency as measured by daily sleep diaries; the only significant ISI correlation with PSG was for sleep efficiency and number of awakenings.

Correlation coefficients were also computed, for both samples separately, between the ISI total scores and scores on instruments measuring related, albeit different, constructs such as sleep quality, fatigue, anxiety, depression, and quality of life (Table 4). The ISI total score was significantly correlated with the PSQI total score, $r_{958} = 0.80$, P < 0.05, suggesting good convergent validity. Significant relationships were also found with other constructs including large correlations with measures of anxiety and depression and moderate correlations with different dimensions of fatigue, especially general and physical fatigue. Significant correlations were also obtained with the SF-12, with a stronger association with the Mental than with the Physical Health component.

ISI sensitivity to detect clinical improvements

A subsample of 146 participants from the Clinical sample who received treatment completed the ISI before and after treatment. Change scores were compared with clinical global improvement ratings obtained from an independent assessor blind to treatment conditions (see Table 5). Participants rated as moderately improved obtained an average ISI change score of −8.4 points (95% CI: −7.2, −9.5), and those rated as markedly improved obtained an average change score of −9.9 points (95% CI: −8.7, −11.0) from baseline to post treatment.

ROC analyses were performed to identify the ISI change score that allows optimal sensitivity and specificity to predict participants who displayed moderate or marked improvement when compared to slight improvement (reference category). These analyses (see Figure 3) revealed that a reduction > 7 points on the ISI was optimal to identify participants with moderate improvements (60% sensitivity, 70% specificity), while an ISI reduction > 8 points was optimal to identify participants with marked improvements (64% sensitivity, 80% specificity).

DISCUSSION

These findings provide additional information about the psychometric properties of the ISI in clinical and population-based samples. Further evidence was obtained about the internal consistency, item response pattern, and convergent validity. New evidence about optimal sensitivity and specificity indices were derived for case finding and for assessing minimally important changes following treatment.

IRT analyses provide new information about the item response pattern in clinical patients. Endorsement of response options 2 to 4 (suggesting more severe insomnia and greater impact) was particularly high for sleep maintenance insomnia symptoms (Items 2 and 3), as well as for satisfaction (Item 4), interference (Item 5), and worry items (Item 7). Conversely, choices 0 and 1 were used infrequently for these items. Except for sleep onset insomnia (Item 1), and to a lesser extent for early morning awakening (Item 3), all other items showed adequate (Item 5) or excellent (Items 2, 4, 6 and 7) discriminative capacity. These latter findings suggest that ratings of sleep onset
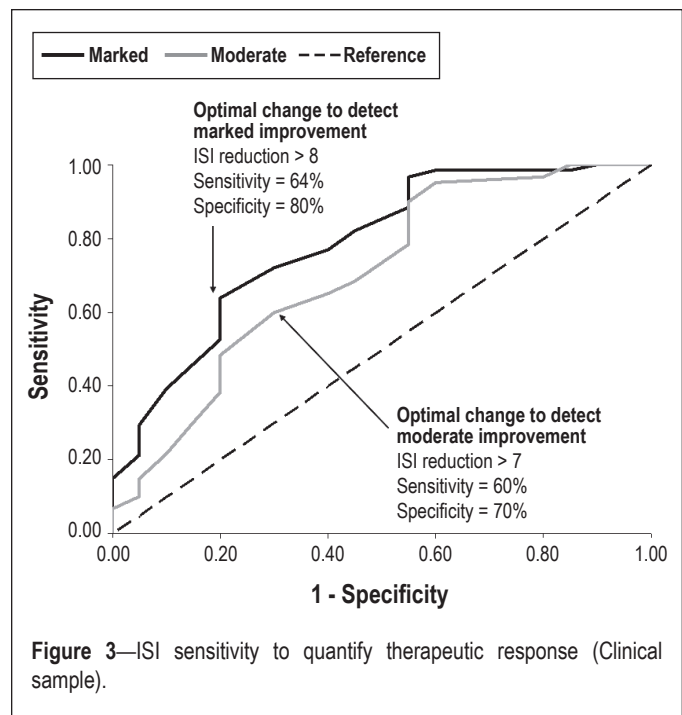
insomnia and early morning awakening may not contribute significantly to the overall insomnia severity, at least in the current clinical sample composed of middle-aged individuals predominantly with primary insomnia. Given this is the first study to report IRT analyses, additional research is warranted with other clinical samples to examine further item response patterns and their relationships with other measures of the insomnia severity construct. Using a more molecular, item-level scrutiny, as opposed to a whole-instrument approach, may also prove useful to identify the best 1 or 2 items for case finding studies.

The results show that different cutoff scores may be useful depending on the specific ISI utilization and research questions. For example, a lower cutoff score may be preferable for epidemiological research aiming to identify insomnia cases in the general population in order to estimate prevalence and incidence rates and burden of disease. However, a cutoff score lower than eight is likely to yield too many false positives, whereas one above 14 would be too stringent and produce too many false negatives. A cutoff score of 10 appears to be the best compromise to achieve optimal balance between sensitivity and specificity in a population-based sample. Likewise, in clinical trials, a cut score of 15 or larger may be too stringent for enrolling patients and actually yield too many false negatives. A cut score of 11 would seem to be the best compromise in this context.

Investigators have traditionally used change scores on sleep latency, wake after sleep onset, and total sleep time to assess the statistical significance of treatment outcome.[15,16,37] Although these sleep parameters represent an important benchmark to assess outcome, when used alone, such end points may not be sufficient to evaluate the clinical significance of outcome.[16] The ISI offers an additional metric that provides complementary information to determine the magnitude and clinical importance of changes in insomnia symptoms. The present data suggest that a change score greater than 7 would be the minimally important difference (MID) to be considered moderately improved (or in partial remission) by an independent rater, whereas a change score of 9 or larger would correspond to a marked improvement rating. These change scores, while informative, may need further validation against additional external criteria. For instance, one recent study proposed a slightly less stringent criterion (i.e., 6-point score reduction) as MID to detect changes on measures of health-related quality of life, work limitations, and fatigue.[19] Additional research is warranted to further document optimal cutoff points as a function of different external criteria and patient populations.

Total ISI scores were positively correlated with subjective sleep estimates and with PSQI total scores, indicating good convergent validity, at least with other patient-reported measures of sleep. Moderate to large correlations with measures of depression, anxiety, and fatigue symptoms and perceived health quality indicated that the ISI measures a construct that is related to or overlapping with these other constructs. Such findings are consistent with several studies showing strong associations between insomnia and psychological symptoms, fatigue, and poorer perceived health.[5,23,38,39] Further research using factor analysis would be needed to study the shared measurement variance with these other dimensions, particularly with anxiety and depression.

Some limitations call for caution in the interpretation of the present findings. First, there was no objective measure or clinical interview to validate the insomnia status in the Community

Figure 3—ISI sensitivity to quantify therapeutic response (Clinical sample).

sample as there was with the Clinical Sample. Results from the Community Sample, which were based on a single yes/no question assessing the participants' perception of having or not a sleep problem, should be interpreted with greater caution than those based on an external standard. Given the nonspecific nature of that question, it is plausible that individuals with sleep problems other than insomnia may have endorsed that item and introduced some confound in the results. This possibility was attenuated, however, by the exclusion of participants who had received a diagnosis of a sleep disorder other than insomnia. Further studies using more stringent "gold standards" would be necessary as evidence supporting our recommendations about optimal ISI cut points with community samples. Secondly, the ability of the ISI to discriminate between primary insomnia and insomnia comorbid with other psychiatric or medical disorders remains unknown. Although this is not the intended use of the ISI, it would be of interest to examine insomnia severity in the context of a more diverse group of patients with both primary and comorbid insomnias.[40] The present findings were derived from community and clinical samples solicited for research participation; it is unclear whether these same cut points would generalize to patients seeking treatment in primary care and could be used to guide the decision to initiate treatment with these patients.

Despite these limitations, brief screening tools such as the Insomnia Severity Index can be useful to epidemiologists for identifying insomnia cases and documenting prevalence and burden of disease. Likewise, the ISI can assist clinicians in their initial evaluation of patients and determine the need for treatment and, once treatment is initiated, in evaluating treatment response. For researchers, the ISI provides a reliable and valid patient-reported outcome to complement specific sleep parameters (e.g., sleep latency, wake after sleep onset) traditionally used as primary end point to support claims of treatment effectiveness.

The measurement of insomnia remains a challenge for clinicians and investigators. Despite the large number of assessment

instruments available in the field, most of them do not meet the specific needs of investigators conducting epidemiological or clinical trial studies. While the present findings add to the psychometric database of the ISI, additional research is needed to refine further this instrument. The development of new instruments is also warranted, perhaps with more focused objectives. Rather than expecting a single instrument to perform well in all domains, it may be more realistic to develop instruments with more specific aims. For instance, an instrument with only two or three key items may be adequate for case findings in large prevalence/incident epidemiological studies. Likewise, with the upcoming DSM-V and expanded focus on dimensional assessment, instruments assessing insomnia symptoms along several dimensions (severity, duration, impact) may prove most helpful to assist clinicians in making a diagnosis. Finally, the design and validation of self-report (patient) measures that can complement sleep diaries or clinician-administered rating scales would prove most valuable in the context of measurement-based patient care.

DISCLOSURE STATEMENT

REFERENCES

1. Morin CM, LeBlanc M, Daley M, Gregoire JP, Merette C. Epidemiology of insomnia: Prevalence, self-help treatments, consultations, and determinants of help-seeking behaviors. Sleep Med 2006;7:123-30.
2. Ohayon MM. Epidemiology of insomnia: What we know and what we still need to learn. Sleep Med Rev 2002;6:97-111.
3. Breslau N, Roth T, Rosenthal L, Andreski P. Sleep disturbance and psychiatric disorders: A longitudinal epidemiological study of young adults. Biol Psychiatry 1996;39:411-8.
4. Daley M, Morin CM, Leblanc M, Gregoire JP, Savard J. The economic burden of insomnia: Direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms and good sleepers. Sleep 2009;32:55-64.
5. Ford DE, Kamerow DB. Epidemiologic study of sleep disturbances and psychiatric disorders. An opportunity for prevention? JAMA 1989;262:1479-84.
6. Ozminkowski RJ, Wang S, Walsh JK. The direct and indirect costs of untreated insomnia in adults in the United States. Sleep 2007;30:263-73.
7. Simon GE, VonKorff M. Prevalence, burden, and treatment of insomnia in primary care. Am J Psychiatry 1997;154:1417-23.
8. Sateia MJ, Doghramji K, Hauri PJ, Morin CM. Evaluation of chronic insomnia. An American Academy of Sleep Medicine review. Sleep 2000;23:243-308.
9. Schutte-Rodin S, Broch L, Buysse D, Dorsey C, Sateia M. Clinical guideline for the evaluation and management of chronic insomnia in adults. J Clin Sleep Med 2008;4:487-504.
10. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. Sleep 2006;29:1155-73.
11. 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.
12. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193-213.
13. Spielman AJ, Saskin P, Thorpy MJ. Treatment of chronic insomnia by restriction of time in bed. Sleep 1987;10:45-56.
14. 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.
15. Martin JL, Ancoli-Israel S. Assessment and diagnosis of insomnia in non-pharmacological intervention studies. Sleep Med Rev 2002;6:379-406.
16. Morin CM. Measuring outcomes in randomized clinical trials of insomnia treatments. Sleep Med Rev 2003;7:263-79.
17. 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.
18. Savard MH, Savard J, Simard S, Ivers H. Empirical validation of the Insomnia Severity Index in cancer patients. Psychooncology 2005;14:429-41.
19. Yang M, Morin CM, Schaefer K, Wallenstein GV. Interpreting score differences in the Insomnia Severity Index: using health-related outcomes to define the minimally important difference. Curr Med Res Opin 2009;25:2487-94.
20. Blais FC, Gendron L, Mimeault V, Morin CM. Assessment of insomnia: Validation of three questionnaires. Encephale 1997;23:447-53.
21. Morin CM, Vallieres A, Guay B, et al. Cognitive behavioral therapy, singly and combined with medication, for persistent insomnia: a randomized controlled trial. JAMA 2009;301:2005-15.
22. Morin CM, Belanger L, LeBlanc M, et al. The natural history of insomnia: A population-based 3-year longitudinal study. Arch Intern Med 2009;169:447-53.
23. LeBlanc M, Beaulieu-Bonneau S, Merette C, Savard J, Ivers H, Morin CM. Psychological and health-related quality of life factors associated with insomnia in a population-based sample. J Psychosom Res 2007;63:157-66.
24. Morin CM. Insomnia: Psychological assessment and management. New York: Guilford Press; 1993.
25. Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. J Psychosom Res 2002;53:737-40.
26. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res 1995;39:315-25.
27. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clin Psychol Rev 1988;8:77-100.
28. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: Psychometric properties. J Consult Clin Psychol 1988;56:893-7.
29. Spielberger CD. Manual for the State-Trait Anxiety Inventory (STAI). Palo Alto, CA: Consulting Psychologists Press. 1983.
30. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220-33.
31. Ware JE, Kosinski M, turner-Bowker DM, Gandek B. How to score version 2 of the SF-12 Health Survey. Lincoln, NE: Quality Metric Incorporated, 2002.
32. Rechtschaffen A, Kales A. A manual of standarized terminology, techniques and scoring system for sleep stages in human subjects. Washington DC: US Government Printing Office. National Institute of Health Publication; 1968.
33. First MB, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM-IV Axis I Disorders: Patient Edition (SCID-I/P, Version 2.0). New York, NY: Biometrics Research Department, New York State Psychiatric Institute; 1997.
34. Embretson SE, Reise SP. Item response theory for psychologists. Mahwah, NJ: Erlbaum, 2000.
35. Samejima F. The graded response model. In: van der Linden WJ, Hambleton RK, eds. Handbook of modern item response theory. New-York: Springer, 1996.
36. Sheu CF, Chen CT, Su YH, Wang WC. Using SAS PROC NLMIXED to fit item response theory models. Behav Res Methods 2005;37:202-18.
37. Morin CM, Bootzin RR, Buysse DJ, Edinger JD, Espie CA, Lichstein KL. Psychological and behavioral treatment of insomnia: Update of the recent evidence (1998-2004). Sleep 2006;29:1398-414.
38. Roth T. Prevalence, associated risks, and treatment patterns of insomnia. J Clin Psychiatry 2005;66 Suppl 9:10-3.
39. Ohayon MM, Roth T. Place of chronic insomnia in the course of depressive and anxiety disorders. J Psychiatr Res 2003;37:9-15.
40. Sarsour K, Morin CM, Foley K, Kalsekar A, Walsh JK. Association of insomnia severity and comorbid medical and psychiatric disorders in a health plan-based sample: Insomnia severity and comorbidities. Sleep Med 2010;11:69-74.