



Development and Validation of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)

Stacie Hudgens¹ · Andrea Phillips-Beyer² · Louise Newton¹ · Dalma Seboek Kinter³ · Heike Benes^{4,5}

Accepted: 13 October 2020 / Published online: 1 November 2020
© The Author(s) 2020

Abstract

Background and Objective Chronic insomnia has major consequences for daytime functioning, yet no fully validated patient-reported outcome instrument for once-daily assessments is available to measure these consequences. This study describes the development and psychometric evaluation of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ).

Methods The Daytime Insomnia Symptom Scale (DISS), an existing 20-item instrument for assessing daytime functioning, was modified to give an 18-item version of the IDSIQ (IDSIQ-18) based on iterative qualitative interviews with 54 subjects with insomnia and expert input. The construct validity and other psychometric properties of the IDSIQ-18 were analyzed based on an interventional study (NCT03056053) in which subjects with insomnia received zolpidem (5 or 10 mg) daily for 2 weeks and an observational study among subjects with no diagnosis of insomnia (good sleepers). Participants in both studies completed the IDSIQ-18 daily for 2 weeks. Exit interviews were conducted with a sample of subjects who completed the interventional study to elicit concepts defining the experience of insomnia, to assess understanding of the response scales, and to determine meaningful change thresholds. Exploratory factor analysis and Rasch analysis were conducted to further assess the structure and latent model for the scoring of the final IDSIQ instrument. Further psychometric evaluation of the final IDSIQ was then conducted.

Results Subjects in both the interventional study ($N = 114$) and observational study ($N = 103$) were predominantly female (65% for subjects with insomnia and 60% for good sleepers). Mean age was 51 years for subjects with insomnia and 45 years for good sleepers. Subjects in the exit interviews ($N = 41$) demonstrated a good understanding of the IDSIQ-18 response scales. Day 1 mean scores were higher (worse) in subjects with insomnia compared with good sleepers. Based on inter-item correlation, exploratory factor, and Rasch analyses and review of the qualitative data, four items were removed. This yielded the final IDSIQ, with 14 items comprising three domains: Alert/Cognition, Mood, and Sleepiness. The domain structure was determined in a confirmatory factor analysis. Evidence of internal consistency reliability was strong: day 1 Cronbach's alpha was 0.917 for IDSIQ total score and 0.806–0.918 for the domains. Test-retest reliability, assessed for subjects with insomnia with no change on the Patient Global Assessment of Disease Severity scale between day 1 and day 8, was also good (intra-class correlation coefficient 0.856–0.911). Meaningful change thresholds derived for this sample using anchor-based approaches were 20 for IDSIQ total score, 9 for the Alert/Cognition domain, 4 for the Mood domain, and 4 for the Sleepiness domain.

Conclusions These studies, which closely followed Food and Drug Administration Guidance for Industry on patient-reported outcome measures, support use of the IDSIQ as a fit-for-purpose measure for deriving valid and reliable endpoints in insomnia clinical research trials and real-world studies.

Stacie Hudgens, Andrea Phillips-Beyer are Co-primary authors.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40271-020-00474-z>) contains supplementary material, which is available to authorized users.

✉ Dalma Seboek Kinter
dalma.seboek-kinter@idorsia.com

Extended author information available on the last page of the article

1 Introduction

The fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) defines insomnia disorder as a predominant complaint of dissatisfaction with sleep quantity or quality that causes distress or impairs functioning in social, occupational, educational, academic, or other settings [1]. Insomnia disorder has a negative effect

Key Points for Decision Makers

The newly developed and validated Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ) is the first self-reported patient-reported outcome tool to evaluate daytime symptoms in people with insomnia disorder.

With a daily recall period, short completion time, and subject feedback on ease of completion, the IDSIQ is patient friendly and has the potential to be used in confirmatory studies of new drugs for insomnia.

on daytime functioning, which manifests as fatigue, sleepiness, comorbid mood disorders, and cognitive impairments [1–5]. In spite of this, current therapies for insomnia have mostly been examined for their effect on sleep, and few data are available to establish improvements in daytime functioning following insomnia treatment [6]. Importantly, there is a dearth of information on the symptom burden and its impact on patients' lives, which are best assessed by patients themselves.

Regulatory authorities now require that efficacy in clinical trials of medical products for insomnia is captured not only by the treating physicians but also by patients themselves using validated patient-reported outcome (PRO) instruments [7, 8]. PROs also have utility in other forms of clinical research and in the routine assessment of patients with insomnia. However, when we reviewed the available literature during the development of a new medical product for insomnia, it became apparent that existing PRO instruments for assessing daytime functioning had not been developed using Food and Drug Administration (FDA) guidance for industry on PRO measures [7]. The most recognized self-report instrument for assessing the severity of insomnia is the Insomnia Severity Index (ISI). The ISI measures both the night-time and daytime components of insomnia, but only includes one item explicitly assessing daytime functioning [9]. Moreover, its 2-week recall period means that it does not capture day-to-day variability in symptoms.

Our review of the literature identified the Daytime Insomnia Symptom Scale (DISS) [10], the Daytime Consequences of Sleep Questionnaire (DCSQ) [11], the Functional Outcomes of Sleep Questionnaire (FOSQ) [12], the Pittsburgh Insomnia Rating Scale (PIRS) [13], and the Profile of Mood States (POMS) [14, 15] as existing PRO instruments that could be used to assess daytime functioning in insomnia. The review of the instruments included the following aspects: content of the questionnaire, evidence of face and content validity, meaning of the items based on their original wordings, relevance of the recall period, level of patient input in questionnaire development, and if psychometric

evaluation was undertaken (including the population, type of analysis, and determination of clinical significance of changes or differences). It was judged that all five instruments had limitations and that a new or modified PRO instrument was needed. Four of the five instruments were assessed to be unsuitable for further development because: they were not publicly available and their development was not documented (DSCQ); they were judged to have a limited number of relevant items (DCSQ, FOSQ, PIRS) or a surfeit of irrelevant items (POMS); or they were developed for an indication other than insomnia (FOSQ) or to assess mood states not linked to a particular indication (POMS). Another instrument for assessing daytime functioning in insomnia, the Sleep Functional Impact Scale (SFIS), was published after our review of the literature [16].

The DISS is a 20-item instrument, developed at the University of Pittsburgh, that assesses daytime functioning using items grouped into Alert/Cognition, Positive Mood, Negative Mood, and Sleepiness/Fatigue domains [10]. It combines nine items for measuring global vigor and affect [17] with 11 items derived from the Hopkins Symptom Checklist-90 [18], Thayer's Activation-Deactivation Adjective Check List [19], and a chart review of 94 patients with chronic insomnia [20]. The factorial structure of the DISS was established in a prospective study in good sleepers and subjects with primary insomnia, which also included psychometric assessment of the DISS items and their ability to capture treatment benefits relating to improved daytime functioning [10]. However, the DISS was not developed using FDA guidance and had limited patient input during item selection and development of the final questionnaire. Moreover, psychometric evaluation of the DISS was limited: test-retest validity, known-groups validity, and responsiveness were not assessed, and meaningful change thresholds were not identified.

It was decided to use the DISS as the starting point for creating and validating a new PRO instrument for capturing patients' experiences of daytime impairment during chronic insomnia. Psychometric properties of the new instrument were explored using data from a single-arm interventional study in subjects with insomnia and an observational study with good sleepers.

2 Early Development of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)

2.1 Methods

Three sets of interviews (Fig. 1) were conducted by independent outcomes research teams in accordance with national laws and regulations. Institutional review board/

independent ethics committee approval was obtained, and all subjects provided informed consent prior to study enrollment.

The qualitative interviews were conducted with subjects diagnosed with insomnia according to the fourth edition of the DSM (DSM-IV) or DSM-5. Subjects were recruited by a combination of clinical sites and recruitment agencies. All interviews were conducted using semi-structured interview guides by researchers trained in qualitative interview techniques; audio recorded and transcribed; and coded and analyzed using ATLAS.ti (Scientific Software Development GmbH). The qualitative data collected were analyzed using grounded theory [21–23] and by content analysis [24, 25].

One-hour face-to-face interviews were held with subjects with insomnia to elicit concepts relating to their insomnia; to subsequently assess the comprehensiveness of the DISS

items and their relevance to the subjects’ experience of insomnia; and to test subjects’ understanding of the DISS’s item wordings, recall period, and response scale. This first set of interviews continued until saturation of concepts was demonstrated. Further 1-hour face-to-face interviews were subsequently conducted with subjects with insomnia to explore the comprehensiveness, relevance, and understandability of a modified version of the DISS comprising 23 items.

The DISS 23 item questionnaire was further modified based on this qualitative research, as well as on discussion with experts in insomnia research. The number of items was ultimately reduced to 18, and the modified instrument was renamed the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)-18. Following the change from a visual analog scale to a numeric rating scale (NRS), qualitative

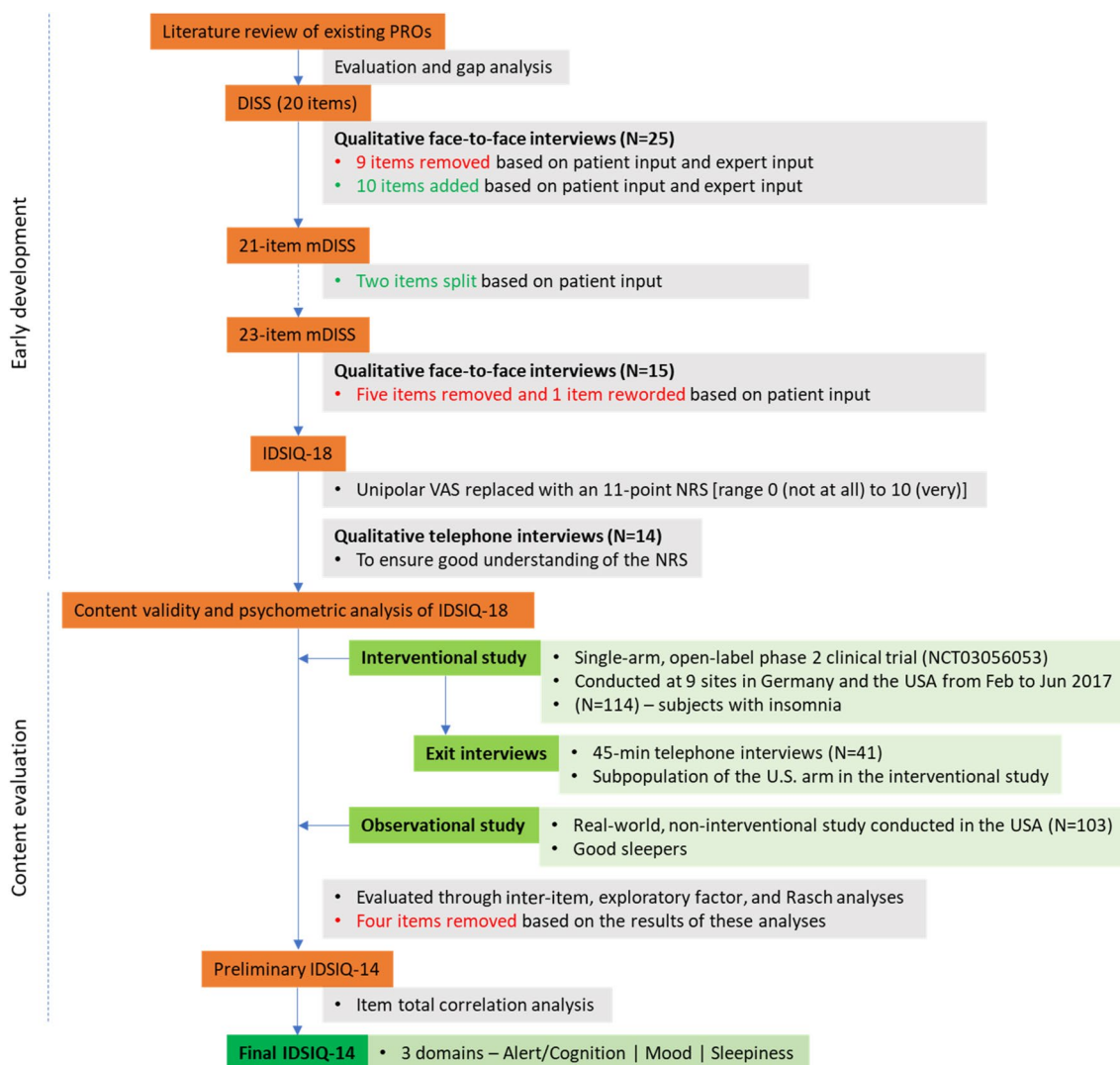


Fig. 1 Development of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ). *DISS* Daytime Insomnia Symptom Scale, *mDISS* modified Daytime Insomnia Symptom Scale, *NRS* numeric rating scale, *PRO* patient-reported outcome, *VAS* visual analog scale

assessment was conducted via telephone interviews with subjects with insomnia to assess their understanding of the IDSIQ-18 response options.

Demographic characteristics of subjects in the qualitative interviews were summarized using descriptive statistics. Semi-structured interview guides available upon request.

2.2 Results

Early development of the IDSIQ is summarized in Fig. 1. Table 1 of the Electronic Supplementary Material (ESM) shows COREQ-required information for the qualitative interviews that contributed to early development of the IDSIQ [26]. Demographic characteristics of the subjects in the qualitative interviews are shown in Table 1.

In the first set of qualitative face-to-face interviews, conducted in 25 subjects diagnosed with insomnia (Table 1), participants generally demonstrated a good understanding of the DISS item wordings. Moreover, there was strong alignment between the item wordings and the language used in the concept elicitation exercise. Based on the interviews, the DISS's response scale was modified to include an absolute zero (i.e., "not at all") point and the assessment time was specified as "today" to capture the variability of symptoms throughout the day (in the original DISS, the recall period was not specified). In addition, four items that were poorly understood or non-specific were removed. Table 2 of the ESM summarizes all item changes during development of the IDSIQ. Further analysis of the qualitative data led to an additional five items being removed and ten items being

Table 1 Demographic characteristics of subjects in the qualitative interviews (full analysis population)

	Subjects with insomnia ^a		
	Face-to-face interviews: 20-item DISS (N = 25)	Face-to-face interviews: 23-item mDISS (N = 15)	Telephone interviews IDSIQ-18 (N = 14)
Age, years			
Mean (SD)	52 (11)	59 (17)	56 (11)
Median (range)	53 (28–71)	65 (33–78)	56.5 (32–76)
Sex, n (%)			
Female	16 (64)	12 (80)	10 (71)
Race, n (%)			
Black or African American	2 (8)	0	3 (21)
Asian	1 (4)	1 (7)	–
White/Caucasian	17 (68)	13 (87)	11 (79)
Other	4 (16)	1 (7)	2 (14)
Unknown	1 (4)	0	–
Ethnicity, n (%)			
Hispanic or Latino	–	–	0
Not Hispanic or Latino	–	–	14 (100)
Not disclosed	–	–	0
Severity of sleep problems, n (%) ^b			
No sleep problems	0	0	1 (7)
Very mild	1 (4)	1 (7)	1 (7)
Mild	4 (16)	2 (13)	1 (7)
Moderate	11 (44)	2 (13)	4 (29)
Severe	7 (28)	10 (67)	6 (43)
Very severe	1 (4)	0	1 (7)
Missing	1 (4)	0	0

DISS Daytime Insomnia Symptom Scale, DSM-IV *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition, DSM-5 *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, IDSIQ Insomnia Daytime Symptoms and Impacts Questionnaire, mDISS modified Daytime Insomnia Symptom Scale, SD standard deviation

^aAccording to DSM-IV (face-to-face interviews) or DSM-5 (telephone interviews)

^bSelf-rated. In the telephone interviews, the recall period was 7 days

added, giving a 21-item modified DISS (mDISS). The iterative methodology used during the concept elicitation supported the view that saturation had been achieved and that the instrument was reflective of the experience of patients with insomnia.

Subsequent revisions, which included the splitting of two items (tired = mentally tired and physically tired; alert = mentally alert and physically alert), yielded a 23-item version of the mDISS. These modifications required additional patient input, and 1-hour face-to-face interviews were conducted with 15 subjects with insomnia (Table 1). It was generally reported that the modified questionnaire was easy to understand. Subjects also demonstrated good understanding of the mDISS items, with the number of subjects who understood each item ranging from 10 to 15. Based on these interviews, five items were removed and one item was reworded. These modifications yielded the IDSIQ-18, which comprised three hypothesized domains: 'Alert/Cognition' (3 items), 'Negative Mood' (8 items), and 'Sleepiness' (7 items). Items 1, 2, 3, 14, and 18 are positively scored (i.e., 0 = worse status; 10 = best status) and require reverse scoring prior to analysis. In addition, based on evidence that an NRS can have superior test-retest validity than a visual analog scale [27], the 10-cm unipolar visual analog scale [10] used in the mDISS was replaced with an 11-point NRS ranging from 0 (not at all) to 10 (very). Subsequent telephone interviews with 14 subjects with insomnia (Table 1) indicated good understanding of the NRS: subjects were able to define the significance of the numbers on the scale in relation to symptom severity.

3 Content and Psychometric Evaluation of the IDSIQ

3.1 Methods

Clinical studies (Fig. 1) were conducted in accordance with International Council for Harmonisation E6 guidelines for Good Clinical Practice, the Declaration of Helsinki, and national laws and regulations. Institutional review board/independent ethics committee approval was obtained for all sites conducting the studies and informed consent was obtained from all subjects prior to study enrollment.

3.1.1 Study Design and Participants

To further evaluate the content validity of the IDSIQ-18 and assess its psychometric properties, data from two studies were analyzed: an interventional study in subjects with insomnia and a non-interventional observational study in

good sleepers. The interventional study was conducted by Idorsia, the sponsor of this research. The observational study and psychometric analyses based on the interventional study were conducted by independent outcomes research teams. In both studies, the IDSIQ-18 was completed as part of an electronic daily diary with mandatory responses for all items.

The interventional study was a single-arm, open-label clinical trial (NCT03056053, EudraCT# 201600425959) conducted at nine sites in Germany and the USA from February to June 2017. Prior to initiation of the interventional study, the IDSIQ-18 was translated into German and reviewed by native German speakers to ensure that the items and instructions were well understood. A total of 114 subjects aged ≥ 18 years who met DSM-5 criteria for insomnia, with self-reported sleep disturbance (≥ 3 nights/week for ≥ 3 months) and an ISI score of ≥ 15 at screening, were enrolled. The design of the interventional study comprised a screening period (14 days) followed by a treatment period (14 days), during which subjects took zolpidem (5 or 10 mg according to the label) once daily in the evening. All subjects completed a daily eDiary in the morning and evening during the screening and treatment periods including the morning of day 15. The IDSIQ-18 was included in the evening diary; the end of treatment IDSIQ-18 assessment was completed in the evening on day 14 (the eDiary was returned during the day on day 15, meaning that no evening IDSIQ-18 assessment was possible). Additional PROs included in the morning and evening diaries were the Patient-Reported Outcomes Measurement Information System (PROMIS) Short Form Sleep Related Impairment (PROMIS Sleep) [28] and Patient Global Assessment of Disease Severity (PGA-S), assessed on days 1, 8, and 15; the ISI [9], assessed on days 1 and 15; and the Patient Global Impression of Change (PGI-C), assessed on days 8 and 15.

The observational study was a multi-site study conducted in the USA to support the validation of the IDSIQ. Its planned duration was February to May 2017. One hundred and three "good sleepers" aged ≥ 18 years without diagnosed insomnia, sleep apnea, or other diagnoses or drug treatments that can lead to sleep problems or fatigue completed the IDSIQ-18 daily for 15 days. All subjects provided written informed consent.

3.1.2 Safety

In the interventional study, adverse events (AEs) were recorded from the signing of informed consent to 30 days after the end of treatment. Treatment-emergent AEs were defined as AEs occurring from the first day of study treatment to the day after the last day of study treatment.

Treatment-emergent AEs were assessed for severity, seriousness, relationship to study treatment, and action taken.

3.1.3 Exit Interviews

At the end of the interventional study, 45-minute exit interviews were conducted by telephone with a subset of subjects to generate a conceptual model of insomnia and the impact of insomnia on daytime functioning and health-related quality of life. Subjects' understanding and interpretation of the IDSIQ-18 response options was also explored. Subjects were asked which three IDSIQ-18 items were most relevant to them, and for these items, were asked to indicate the minimum score change they would consider meaningful. They were also asked how long it took them to complete the IDSIQ-18 and how easy it was to fit into their daily routine. All interviews were held within 2 weeks of completing study treatment.

Interviews were conducted by experienced outcomes researchers using a semi-structured interview guide (available upon request) and were audio recorded and transcribed verbatim. The transcripts were anonymized and were analyzed thematically by grounded theory methods of constant comparison [21–23] and content analysis [24, 25] using ATLAS.ti version 7.0. The conceptual coding dictionary initially comprised codes that were defined *a priori* but was updated iteratively during the coding process.

3.1.4 Statistical Analyses

For both the interventional and observational studies, the following analysis populations were defined: the full analysis population, which comprised all subjects enrolled in the study; the cross-sectional population (CSP), which comprised subjects with non-missing data for the time point being analyzed; and the longitudinal analysis population, which comprised subjects with non-missing data for day 1 and for the subsequent time point being analyzed. The populations used to assess test-retest reliability comprised subjects in the CSP with data for days 1 and 8 and with a ≤ 1 point change on the PGA-S between day 1 and day 8 or no change on the PGA-S between day 1 and day 8. The safety set (interventional study only) comprised all subjects who received at least one dose of zolpidem.

Demographics and baseline characteristics for subjects with insomnia and good sleepers were summarized using descriptive statistics. The percentage of subjects selecting each response option on days 1, 8, and 15 was calculated for each item using the CSP. Floor and ceiling effects were defined as $> 15\%$ of subjects selecting the lowest and highest response options, respectively [29]. Scores for individual IDSIQ-18 items on day 1 were compared between subjects

with insomnia and good sleepers by independent-groups t-test of means for continuous variables. The percentage of subjects who did not complete the IDSIQ-18 was calculated for each day during the treatment period (subjects with insomnia) or study period (good sleepers).

Inter-item correlations were calculated for item scores on day 1 using the CSP. Item pairs with highly correlated responses (> 0.80) were evaluated for redundancy [30–32].

3.1.5 Exploratory Factor Analysis

Using data from day 1, an exploratory factor analysis was performed to evaluate a hypothesized conceptual framework for the IDSIQ-18 comprising three domains: Alert/Cognition (items 1–3), Negative Mood (items 4–11), and Sleepiness (items 12–18) [33]. Scree plots and corresponding eigenvalues were examined to empirically approximate the number of factors for the IDSIQ-18 items. Factor solutions with eigenvalues near to or greater than 1 were examined. Root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) were calculated to evaluate goodness of fit [34]. Models were estimated using weighted least-square mean and variance adjustment and solutions with an oblique rotation (promax) were examined. Items with factor loadings < 0.40 on all factors or that loaded on two or more factors were flagged for further review [35].

3.1.6 Rasch Analysis

The Andrich Rating Scale model [36], which assesses dimensionality, item and response scale fit, and residual correlations between the items (indicating redundancy or overlapping items), was used to confirm the unidimensional factor structure of the IDSIQ items. Criteria for unidimensionality were $> 60\%$ of the variance explained by the Rasch model and $< 5\%$ by the first residual factor. Local dependence (redundancy) was assumed when paired residual correlations were ≥ 0.40 . Acceptable item fit was defined by infit and outfit statistics within the range of 0.5–1.5. To demonstrate the degree to which the items reliably differentiate subjects along the insomnia severity continuum, person and item reliability values from the Rasch model were assessed with a prespecified criterion of ≥ 0.80 (high discrimination).

Four of the IDSIQ-18 items were flagged for removal because of their high inter-item correlations with other items. The deletion of these four items was supported by the results of the exploratory factor and Rasch analyses and a review of the qualitative data specific to these items. The resulting 14-item version of the IDSIQ comprised three domains: Alert/Cognition, Mood, and Sleepiness.

Psychometric evaluation of the IDSIQ-14 was performed using data for subjects with insomnia and good sleepers.

3.1.7 Confirmatory Factor Analysis

Confirmatory factor analysis models were used to confirm the hypothesized three-domain structure of the IDSIQ-14 [37, 38], which was similar to that of the DISS [10]. A full hierarchical model reflecting the relationships of the IDSIQ-14 domains and total score was estimated. The domains were analyzed as first-order factors and IDSIQ-14 total score as a second-order factor. Each domain was also evaluated separately. The confirmatory factor analysis models were estimated by weighted least-square mean and variance-adjusted estimation at day 1 for subjects with insomnia. Model fit was assessed by calculating comparative fit index (CFI: a score ≥ 0.95 indicates good fit), RMSEA (a score < 0.08 indicates acceptable fit), and SRMR (a score < 0.08 indicates good fit) [34, 37, 39]. Adequacy of item fit was evaluated by examining item factor loadings.

3.1.8 Item Total Correlations

Corrected item total Pearson correlations were calculated for the IDSIQ-14, and items for which the correlation with their own domain was < 0.40 were flagged for further review.

3.1.9 Internal Consistency

CSP data from day -6 to day 1 for subjects with insomnia and from day 1 for good sleepers were used to assess the internal consistency of the IDSIQ-14 domains. Cronbach's alpha was calculated for IDSIQ-14 total score and for each domain to assess the homogeneity of the item scores. Cronbach's alpha was recalculated for the domains after removal of individual items. A value > 0.70 indicates acceptable internal consistency [40, 41].

3.1.10 Test-Retest Reliability

Test-retest reliability was assessed by calculating the Shrout and Fleiss intraclass correlation coefficient (ICC) for IDSIQ-14 scores on days 1 and 8: $ICC_{(2,1)} = \text{between-subject mean square} / (\text{between-subject mean square} + \text{within-subject mean square})$. Weekly average IDSIQ-14 scores, calculated as the mean score over days -6 to 1 (day 1) or days 2–8 (day 8), were used in the analyses, where possible (for good sleepers, no IDSIQ-14 data were collected prior to day 1). Subjects with IDSIQ-14 data for fewer than 2 days in a given week were excluded from analyses based on weekly average scores. An ICC value of ≥ 0.70 provides evidence of acceptable test-retest reliability for a scale to be used in detecting group mean differences [42, 43].

3.1.11 Concurrent Validity

Concurrent validity was evaluated at day 1 by calculating Pearson correlation coefficients for IDSIQ-14 total score and domain scores vs ISI total score and PROMIS Sleep T-score. For subjects with insomnia, IDSIQ-14 scores were calculated as the mean score over days -6 to 1. PROMIS Sleep T-score was calculated using a mapping algorithm [28]. A correlation coefficient of > 0.70 indicated a strong correlation; a coefficient of 0.40 – 0.70 a moderate correlation; and a coefficient of < 0.40 a weak correlation.

3.1.12 Known-Groups Validity

Mean weekly average IDSIQ-14 total and domain scores at day 1 were compared between subgroups of subjects with insomnia, grouped according to the ISI total score and PGA-S. Analysis of covariance models were used to assess the significance of differences in IDSIQ-14 total and domain scores, controlling for age and sex. Group means were compared between different ISI and PGA-S severity categories by *post hoc* pairwise testing with Scheffe's test, to adjust for multiple comparisons. Mean weekly average IDSIQ-14 scores were also compared between subjects with insomnia and good sleepers by independent-groups *t* test.

3.1.13 Responsiveness

Responsiveness was calculated using data from the longitudinal analysis population for subjects with insomnia. The final assessment was on day 14 for the IDSIQ-14 and on day 15 for other PROs. For the IDSIQ-14, weekly average scores were calculated as the mean score over days -6 to 1 (day 1), days 2–8 (day 8), and days 9–14 (day 14/15). Differences in mean IDSIQ-14 total score and domain scores between day 1 and days 8 and 14 were calculated for different subgroups of subjects, grouped according to PGI-C, change in PGA-S score, and change in ISI score. The standardized response mean (Cohen's *d*) was calculated by dividing the mean change in the IDSIQ-14 score between day 1 and day 8 or day 14 by the standard deviation of IDSIQ-14 score changes. The magnitude of responsiveness based on the standardized response mean was assessed using Cohen's criteria of $0.20 = \text{small}$, $0.5 = \text{moderate}$, and $0.80 = \text{large}$ [44]. Mean differences in IDSIQ-14 scores between the different PGI-C, PGA-S, and ISI response categories were analyzed using analysis of covariance models adjusted for age, sex, and baseline IDSIQ-14 score. Change categories were compared with the "No change" (PGI-C) or "0" (PGA-S and ISI) category in pairwise fashion by Scheffe's test to adjust for multiple comparisons.

3.1.14 Cumulative Distribution Function and Probability Density Function Curves

Cumulative distribution function and probability density function curves for change in weekly average IDSIQ-14 total score and domain scores according to PGI-C and PGA-S were plotted for subjects with insomnia using the same PGI-C and PGA-S categories as in the analysis of responsiveness.

3.1.15 Discriminant Validity

A multiple linear regression model adjusted for sex and age was used to predict outcomes on the IDSIQ using insomnia status (good sleepers vs subjects with insomnia) as the independent variable. Beta coefficients, standard error, 95% confidence intervals, and *p* values were calculated.

3.1.16 Interpretation of Scores

Thresholds for score changes which subjects might perceive as meaningful were derived by an anchor-based approach using the responsiveness data [45]. PGI-C and PGA-S were chosen as the anchors based on Pearson correlation coefficients >0.40 for correlations with weekly average IDSIQ-14 scores. Differences in mean weekly average IDSIQ-14 scores between day 1 and day 8 or day 14/15 were calculated for each PGI-C and PGA-S change category. Consistent with FDA recommendations [7], multiple estimates based on the anchor-based approach were compared to identify where they converged around a similar set of values that, together with the standardized response mean values and 95% confidence intervals for IDSIQ-14 score changes, could be used to identify meaningful changes.

3.2 Results

Development of the final 14-item IDSIQ is summarized in Fig. 1. Demographic characteristics of the subjects in the clinical studies are shown in Table 2. Subjects in both studies were predominantly female (65% for subjects with insomnia and 60% for good sleepers), with a mean (standard deviation) age of 51 (12) years for subjects with insomnia and 45 (17) years for good sleepers. A majority of subjects were white and non-Hispanic or Latino. Daytime sleepiness based on PGA-S was mild, very mild, or absent in 75% of good sleepers, compared with 8% of subjects with insomnia. Most subjects with insomnia (75%) had an ISI total score of 15–21, indicating insomnia of moderate severity. Current sleep-related health was good to excellent in 97% of good sleepers and fair in 3%.

In the interventional study, the AE data were in accordance with the known safety profile of zolpidem (Table 3 of the ESM). One unrelated AE led to discontinuation of zolpidem; there were no serious or severe AEs.

3.2.1 Exit Interviews

Exit interviews were conducted with 41 US subjects exiting the interventional study (Table 2). Concepts that emerged during the interviews are summarized in Table 4 of the ESM. Table 1 of the ESM shows COREQ-required information for the exit interviews [26].

The most frequently elicited daytime effects of insomnia were tiredness (*n* = 38) and concentration issues (*n* = 17). Twelve of the 15 physical, cognitive, and emotional impacts that emerged during the concept elicitation exercise are captured by the IDSIQ-18 (the concepts that are not captured are pain, numbness, and sadness/depression).

Item 12 (“Energetic”: *n* = 18) and item 1 (“Mentally Alert”: *n* = 14) were the IDSIQ-18 items that resonated most strongly with subjects. Subjects indicated a good understanding of the response scale and were able to provide a clear interpretation of the meaning of a score of 0, 5, and 10 on the response scale.

The mean minimum score change considered meaningful was 3.2 points for the Alert/Cognition domain (*n* = 15), 3.4 points for the Mood domain (*n* = 11), and 3.5 points for the Sleepiness domain (*n* = 18). At the item level, the mean minimum score change that subjects considered meaningful ranged from 2.0 for item 13 (“Effort”) to 4.5 for items 8 (“Irritable”) and 9 (“Impatient”).

Subjects reported that the IDSIQ-18 was easy to complete and took 5–10 minutes. They also reported that it was easy to fit the questionnaire into their daily schedule.

3.2.2 Item Distribution

Missing data were minimal in the interventional study, with 0–7.0% of subjects with insomnia not completing the IDSIQ-18 on any given day in the treatment period. The completion rate was also high for the good sleepers in the observational study, with 9.7–21.4% of subjects not completing the IDSIQ on any given day in the study period.

Analysis of data from day 1 of zolpidem treatment showed that the subjects with insomnia collectively selected all 11 response options (0–10), with the exception of items 14 (“Refreshed”) and 15 (“Mentally Tired”) (Fig. 1a of the ESM). No floor or ceiling effects were observed. By contrast, good sleepers showed a large floor effect: for each item, over 15% of subjects selected a response option of 0 on day 1 and over 40% of subjects selected a response option of 0

Table 2 Demographic characteristics of subjects in the clinical studies (full analysis population)

	Interventional study in subjects with insomnia ^a		Observational study in good sleepers (N = 103)
	Full sample (N = 114)	Exit interviews (N = 41)	
Age, years			
Mean (SD)	51 (12)	57 (10)	45 (17)
Median (range)	53 (19–74)	57 (33–74)	44 (18–74)
Sex, n (%)			
Female	74 (65)	24 (59)	62 (60)
Race, n (%)			
Black or African American	23 (20)	13 (32)	25 (24)
Asian	0	0	2 (2)
White/Caucasian	91 (80)	28 (68)	70 (68)
Other	0	0	4 (4)
Unknown	0	0	2 (2)
Ethnicity, n (%)			
Hispanic or Latino	10 (9)	5 (12)	10 (10)
Not Hispanic or Latino	104 (91)	36 (88)	88 (85)
Not disclosed	0	0	5 (5)
ISI total score at screening, n (%)			
0–7 (no clinically significant insomnia)	0		
8–14 (subthreshold insomnia)	0		
15–21 (moderate insomnia)	86 (75)		
22–28 (severe insomnia)	28 (25)		
Severity of daytime sleepiness at baseline based on PGA-S, n (%)			
None/very mild/mild	9 (8)		77 (75)
Moderate	40 (35)		22 (21)
Severe	61 (54)		4 (4)
Missing	4 (4)		0
Current sleep-related health, n (%) ^b			
Excellent			28 (27)
Very good			47 (46)
Good			25 (24)
Fair			3 (3)

DSM-IV Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, *DSM-5* Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, *ISI* Insomnia Severity Index, *PGA-S* Patient Global Assessment of Disease Severity, *SD* standard deviation

^aAccording to DSM-5

^bCurrent sleep-related health was assessed by asking participants to rate their current health on a 5-point scale, severity of sleepiness over the previous 3 months on a 6-point scale, and average sleepiness in the previous week on a 10-point scale

to 2 (Fig. 1b of the ESM). However, good sleepers selected a response option of at least “8” for each of the 18 items and the “10” response option for 13 of the 18 items. Mean scores for each of the 18 items on day 1 were higher (worse) in subjects with insomnia compared with good sleepers (Table 3).

3.2.3 Inter-Item Correlations

The number of item pairs with correlations > 0.80 on day 1 (which is indicative of redundancy) was 15 for subjects with insomnia and 2 for good sleepers (Table 4). For subjects

with insomnia, item 1 (“Mentally Alert”) was strongly correlated with the two other items in the Alert/Cognition domain: item 2 (“Clear Headed”) and item 3 (“Concentrate”). Moreover, item 6 (“Nervous”) was strongly correlated with another item in the Mood domain: item 11 (“Overly Sensitive”). Items 9 (“Impatient”) and 11 were strongly correlated with item 8 (“Irritable”); item 9 was also strongly correlated with item 10 (“Stressed”). Finally, item 12 (“Energetic”) was strongly correlated with three other items: 1, 14 (“refreshed”), and 18 (“awake”).

3.2.4 Exploratory Factor Analysis

Calculated eigenvalues suggested the presence of two to three separate factors, with good evidence for the first factor (eigenvalue 12.093) and second factor (1.441), and some evidence for a third factor (0.954). In the three-factor model, four items did not conform to the hypothesized domain structure of the IDSIQ-18: item 7 (“Frustrated”) loaded on the Sleepiness domain rather than on the Negative Mood domain and items 12 (“Energetic”), 14 (“Refreshed”), and 18 (“Awake”) loaded on the Alert/Cognition domain rather than the Sleepiness domain (Table 5).

3.2.5 Rasch Analysis

Rasch principal component analysis revealed that 66.6% of the variance on day 1 was explained by the Rasch single-factor dimension. The first residual factor explained 8.8% of the variance, indicating that the IDSIQ is multidimensional (rather than unidimensional). Two item pairs, item 1-item 2 (“Mentally Alert”-“Clear Headed”: 0.74) and item 1-item 3 (“Mentally Alert”- “Concentrate”: 0.65), displayed high residual correlations, suggesting redundancy or local dependence. The outfit statistic of 1.7 for item 4 (“Forgetful”) was slightly above the acceptable range of 0.5–1.5, indicating variability in the item responses at the tails of the severity distribution; however, the infit statistic of 1.5 was within the acceptable range. Person and item reliability values (person = 0.96; item = 0.92) were above the threshold of 0.8.

3.2.6 Generation of the Final IDSIQ Instrument

Based on the results of the inter-item correlation, exploratory factor, and Rasch analyses, as well as the qualitative interviews, four of the IDSIQ-18 items were removed: 1

Table 3 Mean day 1 item scores in subjects with insomnia and good sleepers (full analysis population)

IDSIQ-18 item ^a	Subjects with insomnia (<i>N</i> = 114) Mean (SD)	Good sleepers (<i>N</i> = 89)	<i>T</i> value	<i>P</i> value ^b
Item 1. Mentally Alert	5.5 (2.1)	2.0 (1.9)	12.65	< 0.0001
Item 2. Clear-headed	5.4 (2.0)	1.7 (1.7)	13.56	< 0.0001
Item 3. Concentrate	5.4 (2.2)	1.9 (1.8)	12.16	< 0.0001
Item 4. Forgetful	5.0 (2.6)	2.8 (2.9)	5.71	< 0.0001
Item 5. Worried	5.2 (2.5)	2.4 (2.5)	7.83	< 0.0001
Item 6. Nervous	4.9 (2.4)	1.9 (2.2)	9.17	< 0.0001
Item 7. Frustrated	6.3 (2.2)	2.7 (2.7)	10.17	< 0.0001
Item 8. Irritable	5.5 (2.5)	2.2 (2.4)	9.66	< 0.0001
Item 9. Impatient	5.5 (2.4)	2.2 (2.4)	9.64	< 0.0001
Item 10. Stressed	5.9 (2.3)	2.7 (2.7)	8.97	< 0.0001
Item 11. Overly Sensitive	5.1 (2.6)	1.4 (1.8)	11.97	< 0.0001
Item 12. Energetic	6.0 (1.9)	2.8 (2.2)	10.76	< 0.0001
Item 13. Effort	6.0 (2.1)	3.2 (2.7)	8.08	< 0.0001
Item 14. Refreshed	6.3 (1.9)	3.3 (2.6)	9.04	< 0.0001
Item 15. Mentally Tired	6.1 (2.0)	3.0 (2.5)	9.54	< 0.0001
Item 16. Physically Tired	6.3 (1.9)	2.9 (2.5)	10.76	< 0.0001
Item 17. Sleepy	6.4 (1.9)	2.9 (2.5)	10.65	< 0.0001
Item 18. Awake	6.0 (2.0)	2.5 (2.2)	11.93	< 0.0001

IDSIQ Insomnia Daytime Symptoms and Impacts Questionnaire, *SD* standard deviation

^aItems 1, 2, 3, 14, and 18 were reverse scored prior to analysis so that for all items a higher score indicates a more severe impact

^bIndependent-groups t-test (subjects with insomnia vs good sleepers)

Table 4 Inter-item correlations for the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)-18 on day 1 in subjects with insomnia and good sleepers (cross-sectional population)

Subjects with insomnia			Good sleepers		
Item A	Item B	Pearson correlation coefficient	Item A	Item B	Pearson correlation coefficient
1 Mentally Alert	2 Clear-Headed	0.922	1 Mentally Alert	3 Concentrate	0.869
1 Mentally Alert	3 Concentrate	0.893	2 Clear-Headed	3 Concentrate	0.859
1 Mentally Alert	12 Energetic	0.804			
1 Mentally Alert	18 Awake	0.842			
2 Clear-Headed	3 Concentrate	0.877			
2 Clear-Headed	18 Awake	0.817			
3 Concentrate	18 Awake	0.810			
6 Nervous	11 Overly Sensitive	0.818			
8 Irritable	9 Impatient	0.902			
8 Irritable	11 Overly Sensitive	0.841			
9 Impatient	10 Stressed	0.819			
9 Impatient	11 Overly Sensitive	0.800			
12 Energetic	14 Refreshed	0.823			
12 Energetic	18 Awake	0.855			
14 Refreshed	18 Awake	0.834			

Correlation coefficients were calculated for all item pairs but are only shown for item pairs where the correlation coefficient on day 1 was > 0.8

(“Mentally Alert”), 6 (“Nervous”), 9 (“Impatient”), and 11 (“Overly Sensitive”). The remaining 14 items were renumbered and loaded onto three domains: Alert/Cognition (six items), Negative Mood (four items), and Sleepiness (four items) (Table 6). Although these domains resembled the domains specified in the theoretical framework for the 18-item IDSIQ, the 14 IDSIQ items loaded differently to the hypothesized structure. The “Negative Mood” domain was renamed the “Mood” domain. The final IDSIQ instrument is shown in Fig. 2 of the ESM.

3.2.7 Confirmatory Factor Analysis

Model fit for the full model was acceptable based on CFI (0.950) and SRMR (0.050), but not RMSEA (0.101) (Table 5 of the ESM). The Mood and Sleepiness domains satisfied the prespecified criteria for CFI, RMSEA, and SRMR, indicating that the unidimensional models for these domains had good fit and explained the data well. For Alert/Cognition, although CFI (0.935) and RMSEA (0.201) did not satisfy the prespecified acceptability criteria, good model fit was indicated by SRMR (0.034) and strong factor loadings (range 0.680–0.926).

3.2.8 Item Total Correlation Analysis

For subjects with insomnia, corrected item total correlations for individual IDSIQ-14 items with their own domain were all higher than the threshold of 0.40 on day 1 (range

0.669–0.886). Some items had a weaker correlation for their own domain than for other domains, although in most cases the differences were negligible. Only the item “Energetic” showed a markedly stronger correlation for another domain compared with its own domain: 0.846 for Alert/Cognition vs 0.697 for Sleepiness. The same item also had a stronger correlation with Alert/Cognition than with its own domain among good sleepers (0.614 vs 0.524). The item “Forgetful” had a correlation of 0.342 with its own domain (Alert/Cognition) among good sleepers. All other item total correlations for individual items with their own domain were > 0.40 in good sleepers.

3.2.9 Internal Consistency

Cronbach’s alpha coefficients for IDSIQ-14 total score and the three IDSIQ-14 domains were all above 0.70 (Table 6), indicating acceptable internal consistency.

3.2.10 Test-Retest Reliability

For subjects with no change between day 1 and day 8 on the PGA-S, ICC values for IDSIQ-14 total score and the three IDSIQ-14 domains were all above 0.70 in subjects with insomnia (Table 7), indicating acceptable test-retest reliability. When the analysis was repeated with inclusion of subjects with a 1-point change on the PGA-S, ICC values for IDSIQ-14 total score and the Alert/Cognition and Mood domains remained above 0.70.

Table 5 Exploratory factor analysis of the hypothesized Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ-18) domain structure using day 1 data from subjects with insomnia (full analysis population, $N = 114$)

	Three-factor model		
	Factor 1	Factor 2	Factor 3
Standardized factor loadings			
1. Mentally Alert	0.852	0.022	0.137
2. Clear-Headed	0.876	-0.126	0.253
3. Concentrate	0.877	0.027	0.048
4. Forgetful	0.315	-0.024	0.533
5. Worried	-0.019	0.206	0.637
6. Nervous	0.124	-0.012	0.796
7. Frustrated	-0.010	0.557	0.292
8. Irritable	0.008	0.090	0.878
9. Impatient	0.030	0.169	0.793
10. Stressed	-0.025	0.307	0.649
11. Overly Sensitive	0.085	0.014	0.821
12. Energetic	0.627	0.384	-0.063
13. Activity Effort	0.306	0.495	0.110
14. Refreshed	0.570	0.439	-0.040
15. Mentally Tired	0.148	0.603	0.135
16. Physically Tired	-0.060	0.843	0.108
17. Sleepy	0.077	0.746	0.100
18. Awake	0.671	0.343	-0.033
Model fit statistics			
Chi-square (<i>df</i>)	242.186 (102)		
<i>P</i> value	0.000		
RMSEA	0.110		
90% CI	0.092, 0.128		
Test of close fit: <i>P</i> value	0.000		
SRMR	0.028		

Acceptable factor loadings are shown in bold

CI confidence interval, *df* degrees of freedom, *RMSEA* root mean square error of approximation, *SRMR* standardized root mean square residual

3.2.11 Concurrent Validity

IDSIQ-14 total score and domain scores showed moderate to strong correlations with PROMIS Sleep *T* score in subjects with insomnia (range 0.666–0.778) and good sleepers (0.493–0.755) (Table 8). Correlations with ISI total score were weak in subjects with insomnia (range 0.283–0.315) and weak to moderate in good sleepers (0.321–0.586).

3.2.12 Known-Groups Validity

With the exception of the comparison of the subthreshold and moderate ISI severity categories for the IDSIQ-14 Alert/

Cognition domain ($p = 0.1432$), all pairwise comparisons of different ISI and PGA-S severity categories showed significantly higher mean weekly average IDSIQ-14 total scores and domain scores for the more severe category ($p < 0.05$) (Table 6 of the ESM). Mean weekly average IDSIQ-14 total scores and domain scores were also higher for subjects with insomnia than for good sleepers.

3.2.13 Responsiveness

In an analysis based on data from subjects with insomnia in the interventional study, the IDSIQ-14 demonstrated change over time, with significant decreases in mean weekly average IDSIQ-14 total and domain scores between day 1 and days 8 and 14/15 in subjects who reported a decrease in disease severity based on PGI-C, PGA-S, or ISI (Table 7 of the ESM). However, significant decreases in mean weekly average IDSIQ-14 total and domain scores were also observed in subjects with no change in disease severity based on PGI-C, PGA-S, or ISI.

3.2.14 Cumulative Distribution Function and Probability Density Function Curves

Cumulative distribution function curves showing changes in IDSIQ-14 total score and domain scores according to changes in PGA-S score and PGI-C category indicate greater decreases (improvements) in IDSIQ-14 scores in subjects with insomnia who reported the greatest improvements in symptom severity (Fig. 2). The corresponding probability density function curves are shown in Fig. 3 of the ESM.

3.2.15 Discriminant Validity

In a sex- and age-adjusted analysis, having insomnia was associated with an IDSIQ-14 total score on day 1 that was 46.08 points higher ($p < 0.0001$) compared with good sleepers (Table 9). For individual IDSIQ-14 domains, the corresponding differences were 19.25 points for Alert/Cognition ($p < 0.0001$), 13.26 points for Mood ($p < 0.0001$), and 13.58 points for Sleepiness ($p < 0.0001$).

3.2.16 Interpretation of Scores

Based on anchor-based analyses (Table 7 of the ESM), a 20-point change in the IDSIQ-14 total score was considered meaningful. The corresponding meaningful score changes for the IDSIQ-14 domains were 9 points for the Alert/Cognition domain, 4 points for the Mood domain, and 4 points for the Sleepiness domain.

Table 6 Internal consistency of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)-14 at day 1 (cross-sectional population)

	Cronbach's alpha							
	Subjects with insomnia							Good sleepers ^a
	Day -6 (N = 110)	Day -5 (N = 112)	Day -4 (N = 111)	Day -3 (N = 111)	Day -2 (N = 109)	Day -1 (N = 112)	Day 1 (N = 114)	
IDSIQ-14 total score	0.957	0.953	0.961	0.961	0.964	0.965	0.959	0.917
Alert/Cognition domain	0.934	0.923	0.922	0.935	0.937	0.940	0.933	0.806
Item 1 Clear-Headed	0.915	0.902	0.893	0.917	0.918	0.923	0.910	0.759
Item 2 Concentrate	0.910	0.894	0.890	0.914	0.921	0.916	0.914	0.760
Item 3 Forgetful	0.943	0.931	0.915	0.936	0.938	0.946	0.943	0.838
Item 9 Effort	0.925	0.917	0.914	0.934	0.932	0.929	0.924	0.809
Item 10 Refreshed	0.918	0.906	0.911	0.922	0.925	0.928	0.916	0.740
Item 14 Awake	0.918	0.901	0.921	0.918	0.920	0.930	0.915	0.744
Mood domain	0.903	0.889	0.900	0.915	0.900	0.908	0.897	0.918
Item 4 Worried	0.886	0.869	0.864	0.890	0.894	0.895	0.871	0.919
Item 5 Frustrated	0.877	0.861	0.898	0.902	0.879	0.889	0.889	0.903
Item 6 Irritable	0.858	0.850	0.834	0.876	0.854	0.866	0.849	0.879
Item 7 Stressed	0.876	0.851	0.884	0.889	0.857	0.875	0.858	0.871
Sleepiness domain	0.885	0.849	0.921	0.897	0.915	0.910	0.900	0.860
Item 8 Energetic	0.883	0.869	0.924	0.861	0.894	0.882	0.899	0.888
Item 11 Mentally Tired	0.862	0.785	0.892	0.897	0.899	0.888	0.871	0.788
Item 12 Physically Tired	0.824	0.795	0.876	0.841	0.881	0.896	0.861	0.807
Item 13 Sleepy	0.836	0.778	0.896	0.865	0.885	0.868	0.851	0.787

Cronbach's alpha values for individual items are for the parent domain when the item is removed

^aFor good sleepers, no IDSIQ-14 data were collected prior to day 1

4 Discussion

To the best of our knowledge, this study presents the first self-report PRO instrument developed and validated according to FDA guidelines [7] that can be used to measure the impact of insomnia on daytime functioning. The IDSIQ comprises 14 items with a recall period of “today”. The items are grouped into three domains reflecting daytime effects of insomnia that are commonly encountered in clinical practice [2, 3]: Alert/Cognition, Mood, and Sleepiness. The IDSIQ can be used in clinical trials of insomnia therapies and in observational studies. Following a thorough selection process from among existing instruments, the IDSIQ was rigorously developed using the DISS [10] as the starting point. Its content validity was established in qualitative interviews with 54 subjects and psychometric properties were established and thresholds for clinically meaningful improvements were identified. Given the rigor of its development and adherence to FDA guidelines for PRO development and validation, the IDSIQ can be used for registration studies of insomnia therapies and observational studies to

characterize the deleterious impact of insomnia on daytime functioning.

Confirmatory factor analysis generally supported the domain structure of the final 14-item version of the IDSIQ, although RMSEA indicated suboptimal model fit. However, poor fit for RMSEA is common in smaller samples [46]. Test-retest reliability and known-groups validity of the final IDSIQ were good and concurrent validity based on PROMIS Sleep was acceptable. The poor concurrent validity based on ISI in subjects with insomnia may reflect the fact that the ISI has a 2-week recall period [9], has only one item that directly measures the daytime impact of insomnia, and was developed without patient input. Data from a clinical trial of a sleep medication (zolpidem) were used to assess responsiveness of the IDSIQ to change in subjects with verified DSM-5 insomnia diagnoses. The multiple anchors, anchor threshold criteria, and time points included in the responsiveness analysis made it possible to evaluate whether similar results were obtained across different anchors and different time points. Through this robust approach, responsiveness was found to be acceptable based on decreases in mean IDSIQ-14 scores

Table 7 Test–retest reliability of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)-14 in subjects with insomnia and good sleepers (test–retest populations)

	Subjects with insomnia			Good sleepers				
	N	IDSIQ-14 score, mean (SD)		ICC	N	IDSIQ-14 score, mean (SD)		
		Day 1 ^a	Day 8 ^b			Day 1	Day 8 ^b	
Subjects with no change for PGA-S between day 1 and day 8 ^c	27				44			
IDSIQ-14 total score		70.3 (26.57)	64.4 (25.91)	0.886		34.2 (23.32)	31.8 (20.78)	0.822
Alert/Cognition domain		29.9 (11.24)	27.5 (11.20)	0.880		14.8 (9.50)	13.9 (8.95)	0.750
Mood domain		19.1 (9.21)	17.0 (8.73)	0.911		9.0 (9.76)	7.2 (6.99)	0.765
Sleepiness domain		21.4 (7.36)	19.9 (7.03)	0.856		10.4 (7.38)	10.7 (6.90)	0.731
Subjects with a change of ≤ 1 for PGA-S between day 1 and day 8 ^c	64				73			
IDSIQ-14 total score		78.3 (25.05)	65.1 (22.70)	0.708		35.6 (22.92)	32.3 (20.81)	0.790
Alert/Cognition domain		32.7 (10.84)	27.3 (9.87)	0.703		15.2 (9.74)	14.2 (9.02)	0.742
Mood domain		22.1 (8.27)	17.9 (7.71)	0.751		9.3 (8.95)	7.2 (6.44)	0.705
Sleepiness domain		23.5 (6.96)	19.9 (6.13)	0.664		11.1 (7.75)	10.9 (7.02)	0.706

ICC intra-class correlation coefficient, PGA-S Patient Global Assessment of Disease Severity, SD standard deviation

^aWeekly average score (days –6 to 1)

^bWeekly average score (days 2–8)

^cFor each of the test-retest populations, subjects were required to have IDSIQ data for both day 1 and day 8. Weekly average scores were used for day 1 (days –6 to 1) for subjects with insomnia and for day 8 (days 2–8) for subjects with insomnia and good sleepers

in subjects who reported a decrease in disease severity based on PGI-C, PGA-S, or ISI. The cumulative distribution function and probability density function curves show that a consistent relationship was observed across the continuum of change scores, suggesting that the IDSIQ is sensitive across a range of change values. Finally, despite the limitations of the small sample and dependence on anchor-based methods only, the meaningful change thresholds give a good indication of the levels of improvement that are relevant for this sample of patients with insomnia.

Another instrument, the SFIS, was developed by Bell and colleagues to address the lack of PROs for daytime functioning in insomnia [16]. The SFIS captures the patient experience of how insomnia affects their psychological and cognitive functioning, social activities, and work productivity. Its psychometric properties were validated in a study of 171 subjects with self-reported insomnia and 261 subjects with no history of insomnia. One of the limitations we evaluated during our review of available PRO instruments was whether

the recall period was appropriate. The utility of the SFIS [16] and the DSCQ [11] is limited by a 7-day recall period, which does not capture day-to-day variability in symptoms. In comparison, the IDSIQ offers several advantages such as daily reporting of symptoms, use of an NRS (the SFIS uses a 5-point Likert scale), and calculation of different domain scores (the SFIS offers only a global score), which is advantageous because insomnia is a multi-faceted condition.

Strengths of the presented studies include the capture of insights from a total of 95 subjects with insomnia in interviews during instrument development and validation. Moreover, the subsequent psychometric evaluation was performed in accordance with published FDA guidance for assessing the measurement properties of PROs and for obtaining patient input during clinical development [7, 47]. One limitation is that subjects in the qualitative interviews were naïve to the questionnaire they were asked about, whereas subjects in the exit interviews had completed the IDSIQ-18 every night for 2 weeks. The exit interviews therefore offer a

Table 8 Concurrent validity of the Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)-14 at day 1 in subjects with insomnia and good sleepers (full analysis population)

	IDSIQ-14 total score	IDSIQ-14 Alert/cognition	IDSIQ-14 Mood	IDSIQ-14 Sleepiness	ISI total score	PROMIS Sleep <i>T</i> score
Subjects with insomnia ^a						
IDSIQ-14 total score	1.000	–	–	–	–	–
IDSIQ-14 Alert/ Cognition	0.958***	1.000	–	–	–	–
IDSIQ-14 Mood	0.919***	0.782***	1.000	–	–	–
IDSIQ-14 Sleepiness	0.971***	0.917***	0.863***	1.000	–	–
ISI total score	0.313***	0.283**	0.315***	0.298**	1.000	–
PROMIS Sleep <i>T</i> score	0.774***	0.762***	0.666***	0.778***	0.350***	1.000
Good sleepers						
IDSIQ-14 total score	1.000	–	–	–	–	–
IDSIQ-14 Alert/Cognition	0.894***	1.000	–	–	–	–
IDSIQ-14 Mood	0.802***	0.516***	1.000	–	–	–
IDSIQ-14 Sleepiness	0.895***	0.784***	0.560***	1.000	–	–
ISI total score	0.555***	0.540***	0.321**	0.586***	1.000	–
PROMIS Sleep <i>T</i> score	0.750***	0.703***	0.493***	0.755***	0.767***	1.000

Data are Pearson correlation coefficients

ISI Insomnia Severity Index, PROMIS Patient-Reported Outcomes Measurement Information System, ** $p < 0.01$; *** $p < 0.001$

^aFor subjects with insomnia, weekly average scores (day –6 to day 1) for IDSIQ-14 total score and domain scores were used

perspective that is not directly comparable with the feedback from the qualitative interviews. Another limitation is the fact that no dedicated instrument for assessing daytime functioning was included in the analysis of concurrent validity. Additionally, the validity of the IDSIQ in people with milder insomnia remains to be established, as subjects participating in the interventional study were required to have an ISI score of ≥ 15 . Data from a larger sample could provide further support for the meaningful change thresholds. Moreover, the IDSIQ has so far only been tested with short-term zolpidem use. Its ability to capture longer term treatment effects remains to be established.

5 Conclusions

Through a rigorous process, the IDSIQ was developed and validated as the first self-report PRO instrument with a diurnal recall period for evaluating daytime symptoms in people with insomnia disorder. The short completion time and subject feedback on ease of completion suggest that the IDSIQ is patient friendly. The good test-retest reliability and acceptable concurrent validity and responsiveness of the IDSIQ indicate that it can be used in registration studies of new insomnia drugs.

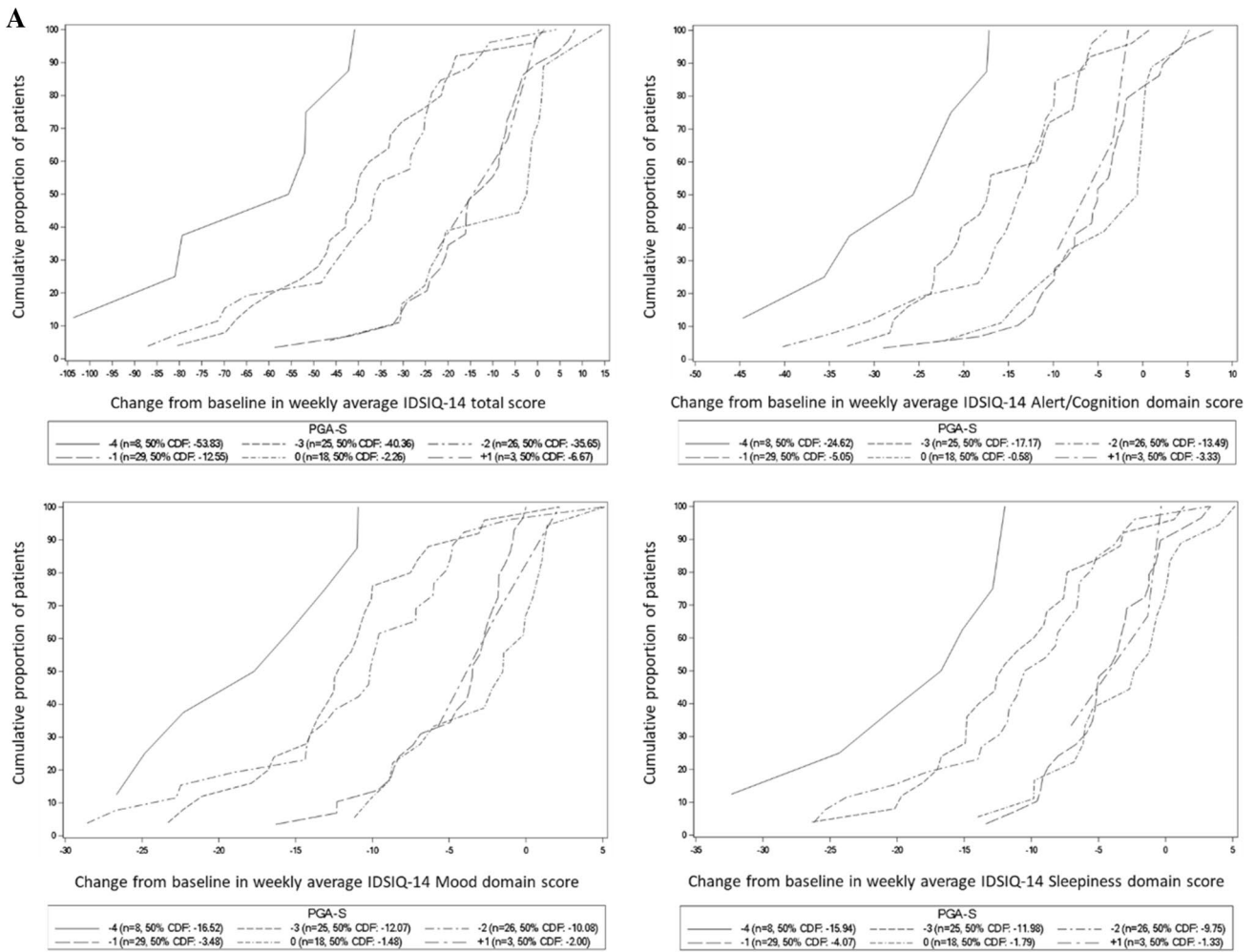


Fig. 2 Cumulative distribution function (CDF) curves for changes in weekly average Insomnia Daytime Symptoms and Impacts (IDSIQ)-14 total score and domain scores according to changes in Patient Global Assessment of Disease Severity (PGA-S) score and Patient

Global Impression of Change (PGI-C) category in subjects with insomnia (full analysis population). **a** Changes in IDSIQ-14 total score and domain scores stratified by change in PGA-S score.

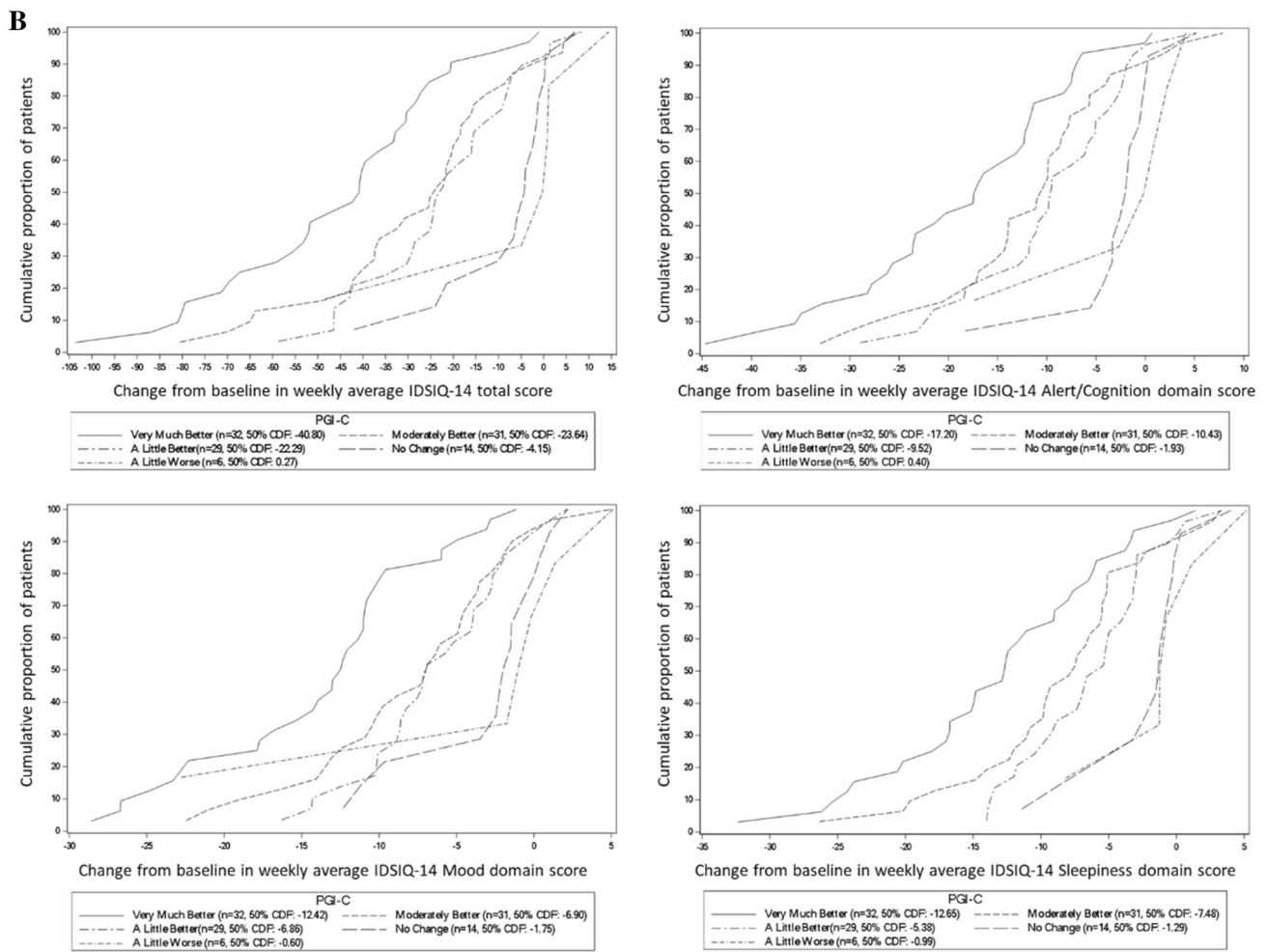


Fig. 2 b Changes in IDSIQ-14 total score and domain scores stratified by PGI-C category. Changes in IDSIQ-14 and PGA-S scores and PGI-C categories are for day 1 to day 14/15

Table 9 Association between insomnia status and Insomnia Daytime Symptoms and Impacts Questionnaire (IDSIQ)-14 scores on day 1 (cross-sectional population)

Domain/variable	β	SE	95% CI	<i>p</i> value
IDSIQ-14 total score				
Insomnia vs good sleepers	46.08	3.40	39.38 to 52.78	< 0.0001
Sex	0.85	3.45	-5.96 to 7.66	0.8060
Age	-0.39	0.12	-0.61 to -0.16	0.0010
Alert/Cognition domain				
Insomnia vs good sleepers	19.25	1.50	16.29 to 22.21	< 0.0001
Sex	-0.20	1.53	-3.21 to 2.81	0.8937
Age	-0.16	0.05	-0.26 to -0.06	0.0018
Mood domain				
Insomnia vs good sleepers	13.26	1.24	10.81 to 15.70	< 0.0001
Sex	1.33	1.26	-1.16 to 3.82	0.2931
Age	-0.12	0.04	-0.20 to -0.03	0.0070
Sleepiness domain				
Insomnia vs good sleepers	13.58	1.05	11.50 to 15.66	< 0.0001
Sex	-0.28	1.07	-2.39 to 1.84	0.7969
Age	-0.11	0.04	-0.18 to -0.04	0.0028

Multiple linear regression model adjusted for sex and age
CI confidence interval, *SE* standard error

Acknowledgments The authors thank Daniel Buysse and the University of Pittsburgh (holders of copyright on IDSIQ) for academic insights and initial development and authorship of the DISS. They also thank Elke Hunsche and Abdel Hmissi for contributions to the study design and data analyses, and Andrew Yaworsky, Steve Hwang, Shannon Walsh, Kathy Lasch, and Andrea Zeytoonjian for qualitative work on debriefing the DISS and developing the modified DISS to generate the initial IDSIQ. Additionally, the authors thank Bruno Flamion and Randall Watson for critical review of the manuscript. Medical writing support was provided by Stephen Gilliver and Rangan Gupta of Evidera, Inc., and was paid for by Idorsia.

Declarations

Funding This work, and all studies, were funded by Idorsia Pharmaceuticals Ltd. and formerly (prior to June 2017) by Actelion Pharmaceuticals Ltd.

Conflicts of Interest/Competing Interests Stacie Hudgens and Louise Newton are employees of Clinical Outcomes Solutions; Clinical Outcomes Solutions was funded by Actelion and Idorsia to conduct the observational study, psychometric validation, and exit interviews. Andrea Phillips-Beyer is the director of Innovus Consulting Ltd., which provided consulting services for Idorsia Pharmaceuticals Ltd. and received payment for the work outlined in the manuscript. Dalma Seboek Kinter is an employee and shareholder of Idorsia Pharmaceuticals Ltd. Heike Benes is an employee of Somni Bene Institut für Medizinische Forschung und Schlafmedizin Schwerin GmbH and University of Ros-

tock Medical Center, Germany, which received financial support for conducting the interventional study.

Ethics Approval All studies including the qualitative interviews were reviewed and approved by an institutional review board/independent ethics committee and all procedures were in accordance with International Council for Harmonization Good Clinical Practice, the Declaration of Helsinki, and national laws and regulations.

Consent to Participate All study subjects provided written informed consent prior to enrollment.

Consent for Publication All authors have reviewed and approved the manuscript and consent to its publication.

Data Availability and Material The data generated during and/or analyzed during the current study are proprietary and not publicly available but may be available from the corresponding author on reasonable request.

Code Availability Not applicable.

Authors' Contributions SH as an independent PRO lead scientist designed, and conducted the observational study, contributed to the study design of the interventional study, performed psychometric analyses on the interventional and observational studies, interpreted data, and was lead author on the related reports. APB performed psychometric and validation analyses on the interventional and observational studies, performed analyses on qualitative study data, interpreted data, and drafted the manuscript in collaboration with the other authors. LN designed, performed, and interpreted the exit interviews. DSK provided input to the design and data interpretation for the qualitative studies, and for the interventional and observational studies. HB conducted the interventional study and interpreted data. All authors critically reviewed the manuscript and approved the final version of the manuscript for submission.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. 2013. <https://www.psychiatry.org/psychiatrists/practice/dsm/dsm-5>. Accessed 11 Jun 2020.
2. Carey TJ, Moul DE, Pilkonis P, Germain A, Buysse DJ. Focusing on the experience of insomnia. *Behav Sleep Med.* 2005;3(2):73–86. https://doi.org/10.1207/s15402010bsm0302_2.
3. Kyle SD, Espie CA, Morgan K. “Not just a minor thing, it is something major, which stops you from functioning daily”: quality of life and daytime functioning in insomnia. *Behav Sleep Med.*

- 2010;8(3):123–40. <https://doi.org/10.1080/15402002.2010.487450>.
4. Shekleton JA, Flynn-Evans EE, Miller B, Epstein LJ, Kirsch D, Brogna LA, et al. Neurobehavioral performance impairment in insomnia: relationships with self-reported sleep and daytime functioning. *Sleep*. 2014;37(1):107–16. <https://doi.org/10.5665/sleep.3318>.
 5. 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(5):487–504.
 6. Sandlund C, Hetta J, Nilsson GH, Ekstedt M, Westman J. Impact of group treatment for insomnia on daytime symptomatology: analyses from a randomized controlled trial in primary care. *Int J Nurs Stud*. 2018;85:126–35. <https://doi.org/10.1016/j.ijnurstu.2018.05.002>.
 7. US FDA. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009. <https://www.fda.gov/media/77832/download>. Accessed 11 Jun 2020.
 8. EMA. Guideline on medicinal products for the treatment of insomnia. London: European Medicines Agency; 2009. https://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/11/WC500011987.pdf. Accessed 7 Jul 2020.
 9. Bastien CH, Vallieres A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med*. 2001;2(4):297–307. [https://doi.org/10.1016/s1389-9457\(00\)00065-4](https://doi.org/10.1016/s1389-9457(00)00065-4).
 10. Buysse DJ, Thompson W, Scott J, Franzen PL, Germain A, Hall M, et al. Daytime symptoms in primary insomnia: a prospective analysis using ecological momentary assessment. *Sleep Med*. 2007;8(3):198–208. <https://doi.org/10.1016/j.sleep.2006.10.006>.
 11. Hays RD, Morlock RJ, Spritzer K, Drake C, Roth T. Psychometric properties of the Restorative Sleep Questionnaire and Daytime Consequences of Sleep Questionnaire. In: 19th Annual Meeting of the Associated Professional Sleep Societies (APSS); 22 June 2005; Denver (CO).
 12. Weaver TE, Laizner AM, Evans LK, Maislin G, Chugh DK, Lyon K, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*. 1997;20(10):835–43.
 13. Moul DE, Pilkonis PA, Miewald JM, Carey TJ, Buysse DJ. Preliminary study of the test-retest reliability and concurrent validities of the Pittsburgh Insomnia Rating Scale (PIRS). *Sleep*. 2002;25:A246–7.
 14. McNair D, Lorr M, Doppleman L. Manual for the Profile of Mood States. San Diego: Educational and Industrial Testing Service; 1971.
 15. Yeun EJ, Shin-Park KK. Verification of the Profile of Mood States-Brief: cross-cultural analysis. *J Clin Psychol*. 2006;62(9):1173–80. <https://doi.org/10.1002/jclp.20269>.
 16. Bell C, McLeod LD, Nelson LM, Fehnel SE, Zografos LJ, Bowers B. Development and psychometric evaluation of a new patient-reported outcome instrument measuring the functional impact of insomnia. *Qual Life Res*. 2011;20(9):1457–68. <https://doi.org/10.1007/s11136-011-9885-8>.
 17. Monk TH. A visual analogue scale technique to measure global vigor and affect. *Psychiatry Res*. 1989;27(1):89–99. [https://doi.org/10.1016/0165-1781\(89\)90013-9](https://doi.org/10.1016/0165-1781(89)90013-9).
 18. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behav Sci*. 1974;19(1):1–15. <https://doi.org/10.1002/bs.3830190102>.
 19. Thayer RE. Factor analytic and reliability studies on the Activation-Deactivation Adjective Check List. *Psychol Rep*. 1978;42(3 PT 1):747–56. <https://doi.org/10.2466/pr0.1978.42.3.747>.
 20. Moul DE, Nofzinger EA, Pilkonis PA, Houck PR, Miewald JM, Buysse DJ. Symptom reports in severe chronic insomnia. *Sleep*. 2002;25(5):553–63.
 21. Glaser B, Strauss A. The constant comparative methods of qualitative analysis: discovery of grounded theory. New York: Aldine de Gruyter; 1967.
 22. Charmaz K. Grounded theory. In: Smith JA, Harre R, Van Langenhove L, editors. Rethinking methods in psychology. London: Sage; 1995. p. 27–49.
 23. Strauss A, Corbin J. Basics of qualitative research: techniques and procedures for developing grounded theory. London: Sage; 1998.
 24. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15(9):1277–88. <https://doi.org/10.1177/1049732305276687>.
 25. Lasch KE, Marquis P, Vigneux M, Abetz L, Arnould B, Bayliss M, et al. PRO development: rigorous qualitative research as the crucial foundation. *Qual Life Res*. 2010;19(8):1087–96. <https://doi.org/10.1007/s11136-010-9677-6>.
 26. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care*. 2007;19(6):349–57. <https://doi.org/10.1093/intqhc/mzm042>.
 27. Gries K, Berry P, Harrington M, Crescioni M, Patel M, Rudell K, et al. Literature review to assemble the evidence for response scales used in patient-reported outcome measures. *J Patient Rep Outcomes*. 2018;2:41. <https://doi.org/10.1186/s4168-018-0056-3>.
 28. Unpublished manual for the Patient-Reported Outcomes Measurement Information System (PROMIS) Version 1.1. PROMIS Cooperative Group; 2008. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.169.5144&rep=rep1&type=pdf>
 29. Janssen CA, Oude Voshaar MAH, Ten Klooster PM, Jansen T, Vonkeman HE, van de Laar M. A systematic literature review of patient-reported outcome measures used in gout: an evaluation of their content and measurement properties. *Health Qual Life Outcomes*. 2019;17(1):63. <https://doi.org/10.1186/s12955-019-1125-x>.
 30. Carmines E, Zeller R. Reliability and validity assessment. In: Carmines E, Zeller R, editors. Quantitative applications in the social sciences. Beverly Hills: Sage Publications; 1979.
 31. Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. 2013.
 32. Streiner D, Norman G, Cairney J. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press; 2014.
 33. Howard MC. A review of exploratory factor analysis decisions and overview of current practices: what we are doing and how can we improve? *Int J Hum Comput Interact*. 2016;32(1):51–62.
 34. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. Testing structural equation models. Beverly Hills: Sage Publications; 1993. p. 136–62.
 35. Stevens J. Applied multivariate statistics for the social sciences. 2nd ed. Hillsdale: Erlbaum; 1992.
 36. Andrich D. A rating scale formulation for ordered response categories. *Psychometrika*. 1978;43:561–73.
 37. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6:1–55.
 38. Kline RB. Principles and practice of structural equation modeling. 3rd ed. New York, (NY): Guilford Press; 2010.
 39. Hu L, Bentler PM. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol Methods*. 1998;3:424–53.
 40. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297–334.

41. Nunnally JC, Bernstein IH. The assessment of reliability. In: Psychometric theory. New York: McGraw Hill; 1994. p. 248–92.
42. Nunnally JC, Bernstein IH. Psychometric theory. New York: McGraw-Hill; 1994.
43. Leidy NK, Revicki DA, Geneste B. Recommendations for evaluating the validity of quality of life claims for labeling and promotion. *Value Health*. 1999;2(2):113–27. <https://doi.org/10.1046/j.1524-4733.1999.02210.x>.
44. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.
45. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102–9. <https://doi.org/10.1016/j.jclinepi.2007.03.012>.
46. Kenny DA, Kaniskan B, McCoach DB. The performance of RMSEA in models with small degrees of freedom. *Sociol Methods Res*. 2015;44(3):486–507.
47. US FDA. Guidance for industry, Food and Drug Administration staff, and other stakeholders. Patient-focused drug development: methods to identify what is important to patients (draft guidance). 2019. <https://www.fda.gov/media/131230/download>. Accessed 17 Aug 2020.

Affiliations

Stacie Hudgens¹ · Andrea Phillips-Beyer² · Louise Newton¹ · Dalma Seboek Kinter³ · Heike Benes^{4,5}

¹ Clinical Outcomes Solutions, Tucson, AZ, USA

² Innovus Consulting Ltd., London, UK

³ Idorsia Pharmaceuticals Ltd., Allschwil, Switzerland

⁴ Somni Bene Institut für Medizinische Forschung und Schlafmedizin Schwerin GmbH, Schwerin, Germany

⁵ University of Rostock Medical Center, Rostock, Germany