CLINICAL REVIEW

Questionnaires that screen for multiple sleep disorders

Karen J. Klingman [a, *], Carla R. Jungquist [b], Michael L. Perlis [c, d]

[a] Upstate Medical University (SUNY Upstate), College of Nursing, USA
[b] University at Buffalo (SUNY Buffalo), School of Nursing, USA
[c] University of Pennsylvania, Behavioral Sleep Medicine Program, Department of Psychiatry, USA
[d] University of Pennsylvania, School of Nursing, USA

SUMMARY

The goal of this review was to identify, describe, and evaluate the existing multiple sleep disorders screening questionnaires for their comprehensiveness, brevity, and psychometric quality. A systematic review was conducted using Medline/PubMed, cumulative index to nursing & allied health literature, health and psychosocial instruments and the "grey literature". Search terms were "sleep disorders, screening, questionnaires, and psychometrics". The scope of the search was limited to English language articles for adult age groups from 1989 through 2015. Of the n = 2812 articles identified, most were assessment or treatment guideline reviews, topical reviews, and/or empirical articles. Seven of the articles described multiple sleep disorders screening instruments. Of the identified instruments, two questionnaires (the Holland sleep Disorders questionnaire and sleep-50) were evaluated as comprehensive and one questionnaire (the global sleep assessment questionnaire [GSAQ]) was judged to be both comprehensive and efficient. The GSAQ was found to cover four of the six core intrinsic disorders, sleep insufficiency, and daytime sequela with 11 questions. Accordingly, the GSAQ is the most suitable for application as a general sleep disorders screener. Additional work is required to validate this instrument in the context of primary care. Finally, the future development of multiple sleep disorders screening questionnaires should not only cover all six intrinsic sleep disorders but also acquire some basic demographic information (age, sex, body mass index, presence/absence of bed partner, work status and shift) and some limited data regarding sleep sufficiency and the daytime consequences of sleep disturbance.

© 2016 Elsevier Ltd. All rights reserved.

Introduction

Over the course of the last two decades, it has become increasingly clear that undiagnosed and untreated sleep disorders increase the individual's risk for new onset medical and psychiatric illness [1–4]. This, in combination with the high prevalence of sleep disorders in the population (more than 40% by some estimates [2]) underscores the need for the increased assessment of sleep disorders at the primary care level. While an increased exposure to sleep medicine during clinical training (both pre-professionally and as part of continuing professional education) would be an ideal way to address this issue, such curricular changes would be difficult to implement and take years, if not decades, to affect clinical practice. With increasing awareness of a need to know patients' sleep disorder status, primary care providers may remain uncertain as to how to make differential assessments without a sleep medicine consult or at least a general screening tool.

In some ways this scenario parallels that which occurred with depression in the late 1990s and early 2000s. At that time, primary care had been sensitized to the need for assessment and treatment or referral for depression but lacked the tools required to accomplish this end. While several instruments existed to quantify depression severity (e.g., the 30-item inventory for depressive symptomatology – IDS [5,6], the 21-item Beck depression inventory – BDI [7,8], the 24-item Hamilton rating scale for depression – HRSD-24 [9], and the 30-item geriatric depression scale – GDS-30 [10]), none were suitable for primary care practice given their length and focus on severity. This problem was remedied with the development of a nine-item screener, patient health questionnaire [11] (PHQ-9). This instrument was both brief and allowed non-specialists to make a diagnostic assessment. This scenario suggests that the need to assess for sleep disorders within primary care might be remedied in a similar way: a brief screening tool could be used to assess for the incidence of sleep disorders.

* Corresponding author. Tel.: +1 585 208 5336.
E-mail address: klingmak@upstate.edu (K.J. Klingman).

List of abbreviations

AASM	American Academy of Sleep Medicine
ASPS	advanced sleep phase syndrome
AUC	area under the curve
ASQ	Auckland sleep questionnaire
CINAHL	cumulative index to nursing & allied health literature
BDI	Beck depression inventory
BMI	body mass index
COSMIN	consensus-based standards for the selection of health measurement instruments
CRSD	circadian rhythm sleep wake disorders
DSM	diagnostic and statistical manual of mental disorders
DSPS	delayed sleep phase syndrome
ESS	Epworth sleepiness scale
GSAQ	global sleep health assessment questionnaire
GDS	geriatric depression scale
HaPI	health and psychosocial instruments
HRSD	Hamilton rating scale for depression
HSDQ	Holland sleep disorders questionnaire insomnia
ICSD	international classification of sleep disorders
IDS	inventory for depressive symptomatology
INS	insomnia
ISDI	Iowa sleep disturbances inventory
IRLS	international restless legs scale
ISI	insomnia severity index
MeSH	medical subject headings
n/a	not applicable or not available
NAR	narcolepsy
NCSDR	National Center on Sleep Disorders Research
NINDS	National Institute of Neurological Disorders and Stroke
NSF	National Sleep Foundation
OSA	obstructive sleep apnea
PAR	parasomnias
PDF	portable document format
PHQ	patient health questionnaire
PSQI	Pittsburgh sleep quality index
pt	patient
RLS	restless legs syndrome
ROC	receiver operator characteristic curve
SBSM	Society of Behavioral Sleep Medicine
SD	sleep disorders
SDQ	sleep disorders questionnaire
SQAW	sleep questionnaire and assessment of wakefulness
SSC	sleep symptom checklist
STOP	snore, tired, observed, pressure

This prospect, however, is complicated by the need to assess not one illness, but several which encompass the core sleep disorders (or classes of disorders) including insomnia, sleep disordered breathing, circadian rhythm disorders, restless legs syndrome/periodic leg movements of sleep, parasomnias and narcolepsy.

At present, there are a variety of single condition measures available (e.g., the insomnia severity index − ISI [12], the Berlin questionnaire for sleep apnea [13], the STOP (snore, tired, observed, pressure) questionnaire for sleep apnea [14], the International restless legs syndrome rating scale − IRLS [15], and Epworth sleepiness scale − ESS [16]). These instruments, however, were created for use by specialists to assess the severity of specific disorders and are not practical to use in combination for screening. The only well-known, and well-studied, global instrument presently in use is the Pittsburgh sleep quality index (PSQI) [17]. This instrument, while the first of its kind and useful for the quantification of global sleep disturbance severity, was not intended to be used as a stand-alone screening tool by non-sleep professionals. Thus, what is needed is a "one page" instrument that can be completed by patients and provide the primary care clinician with an easy way to discern 1) which presenting complaints map onto specific sleep disorders and 2) whether the sleep complaints are of a severity and/or frequency to warrant treatment and/or referral. The present review was undertaken to identify what instruments are presently available to assess multiple sleep disorders, and which do so in a manner that is suitable for use in primary care.

Methods

This analysis first identified candidate questionnaires by reviewing the literature and then appraised questionnaires for potential usefulness in primary care.

Literature review

A review of the literature was conducted using databases (Medline/PubMed, CINAHL − cumulative index to nursing & allied health literature, and HaPI − health and psychosocial instruments) and "grey literature" (information presented in academic, government, foundation, and industry reports). The database search terms were: a) sleep disorder AND diagnosis, b) sleep disorder AND psychometric*, c) sleep disorder AND questionnaire, d) sleep disorder AND screen*, and (e) sleep disorder AND validation (where * indicates a wildcard to allow for multiple word endings). The search was restricted to English language articles for all adult age groups (ages 18+) from 1989 through 2015. Grey literature was searched for "sleep questionnaire" with both Google and Google Scholar. Additional sources for grey literature were the National Guideline Clearinghouse (search term "sleep disorder") [18], Centers for Disease Control and Prevention National Health and Nutrition Examination Survey website (question items related to sleep) [19], and the American Academy of Sleep Medicine website (search term "sleep disorder") [20]. Date of last search for grey literature was March 25, 2015.

Articles/questionnaires with studies that met the following criteria were included for subsequent feature assessment and appraisal − the screening instrument: 1) was based on self-report, 2) assessed at least three sleep disorders, and 3) was evaluated psychometrically. Articles were excluded if the studies focused primarily on specific diseases, if they were reviews or practice guideline papers, or if they described structured interview procedures. Reference lists from included database articles were also scanned. Following review of this search strategy with a health sciences librarian, MeSH (medical subject headings) terms "sleep disorders", "sleep disorders, intrinsic", "sleep disorders, circadian rhythm", and "dyssomnias" were also searched using PubMed (advanced search builder). The resulting list of candidate questionnaires proposed for feature assessment was then reviewed by two experts in the field of sleep medicine to identify possibly-missed questionnaires.

Quality appraisal

Questionnaires were assessed for their practicality and psychometric soundness. Practicality was assessed in terms of

instrument comprehensiveness and brevity. Psychometric soundness was assessed in terms of the manner in which the instrument was constructed and whether or not the instrument was evaluated for validity. Practicality was preferentially weighted in the scoring strategy. A total of 14 points were possible for each instrument, with six points possible for comprehensiveness, four points possible for brevity, and four points possible for psychometric soundness. Appraisal scores were assigned by consensus of first and third authors.

Practicality

For the comprehensiveness assessment, questionnaires were scored one point for each of the six intrinsic sleep disorders assessed. Brevity was scored on a weighted three point scale: four points for questionnaires with ≤20 items; two points for questionnaires of between 20 and 25 items; and zero points for questionnaires larger than 25 items. The 25 item limit (and the scoring for comprehensiveness) was adopted based on the following rationale. Given that single-disorder screeners utilize up to nine items to assess for one disorder (e.g., the PHQ-9 [11]), it could be argued that up to nine questions per disorder should be allowed for a multi-disorder screener. In the present case this would suggest that a screener for the six intrinsic sleep disorders would be up to 54 items in length. An instrument of this size, however, would be expected to take between 27 and 54 min to complete. Given that patients typically wait about 15 min before a primary care appointment [21], a questionnaire of this length would not be practical. Given that some amount of time must also be set aside for other pre-visit paperwork and/or assessments, a reasonable estimate for the amount of time available would be approximately 10–15 min. This being the case, and working backwards from the time allowed (and given that each item on a questionnaire takes about 30 s to complete [22] [varying with the questionnaire format]), it follows that the maximum number of items for a multi disorder sleep screener would be about 25 items. Given this framework, the scoring system for brevity utilized a threshold of 25 items.

Psychometric soundness

In order to provide a broad assessment of whether questionnaires were based on solid methods of construction and validation, four yes/no questions were used to characterize each instrument.

1) Were all of the items standardized, i.e., adopted from previously established instruments or ratified by recognized expert consensus panels?
2) Was an effort made to streamline the instrument using a statistical technique (such as, confirmatory factor analysis or a logistic regression, for example, where items with low weightings were eliminated)?
3) Was the item cluster (subscale) for each diagnosis shown to be relatively sensitive and specific for the target disorder, i.e., did the authors provide true/false positive rates, sensitivity/specificity rates, and/or ROC/area under the curve (AUC) analyses?
4) Was the instrument evaluated for sensitivity to change in patient status, i.e., spontaneous recovery or treatment-related symptom improvement?

Results

Literature review results

As can be seen in Fig. 1, 2812 articles were reviewed (2398 from databases and 414 from grey literature), 2805 were excluded and seven full-text articles that met the inclusion/exclusion criteria were assessed and reviewed further. The expert consultants identified additional screening tools, but these did not cover three or more sleep disorders and so were not considered in the final review. Results using the additional librarian-suggested MeSH terms with PubMed (advanced search builder) yielded no new inclusion candidates. The seven questionnaires identified were: 1) Auckland sleep questionnaire (ASQ) [23], 2) global sleep assessment questionnaire (GSAQ) [24], 3) Holland sleep disorders questionnaire (HSDQ) [25], 4) ISDI (Iowa sleep disturbances inventory (ISDI) [26], 5) sleep disorders questionnaire (SDQ) [27], 6) sleep-50 [28], and 7) sleep symptom checklist (SSC) [22]. Table 1 summarizes the number of items, response options, number of covered sleep disorders, and time period required for each of the seven questionnaires. Summary information for each of the seven instruments is provided below (instruments are reviewed in alphabetic order).
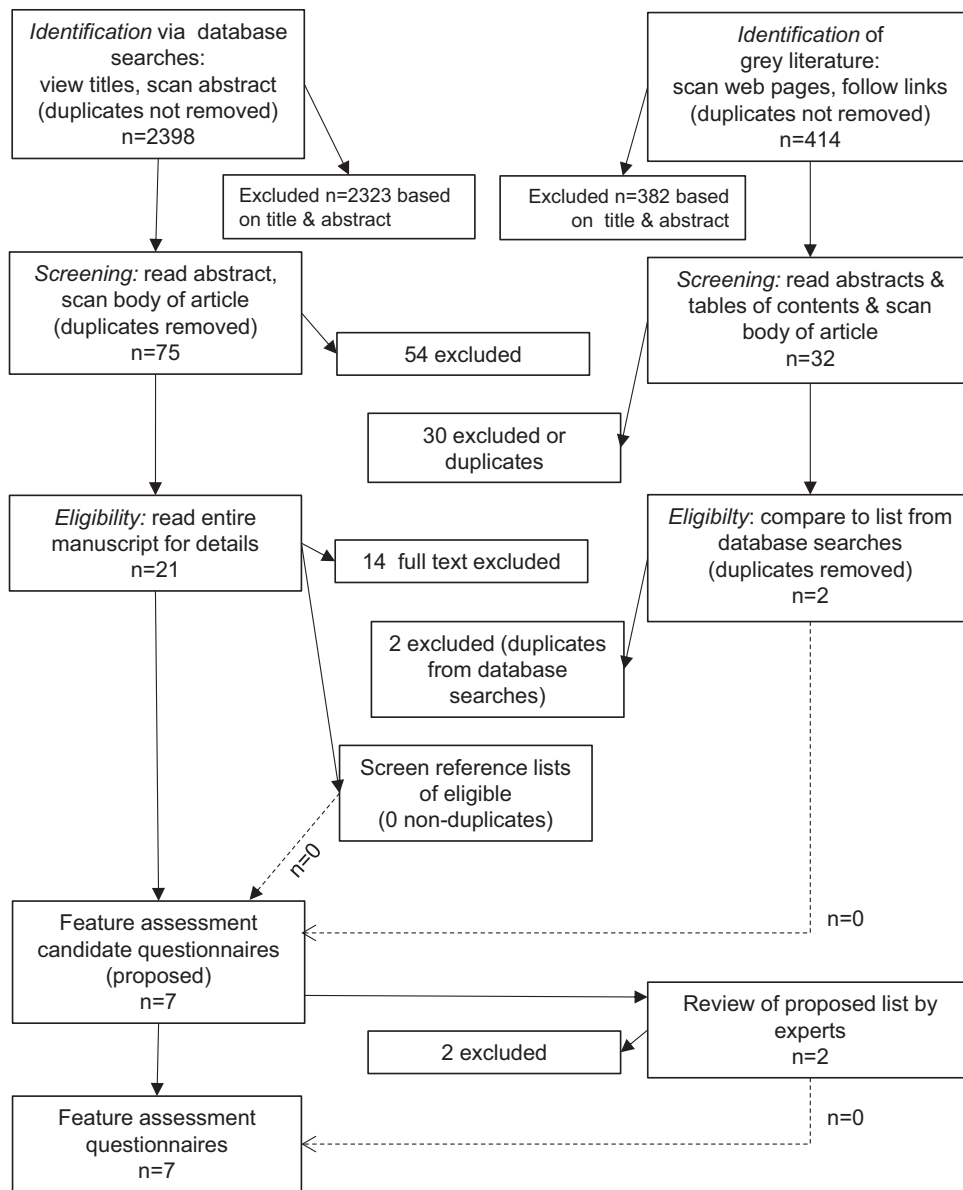
Auckland sleep questionnaire (ASQ)

The ASQ [23] comprises 30 items and is arrayed over six pages. Items included are questions regarding demographics, general health conditions, work schedule, drug and alcohol usage, menopausal status, anxiety and depression, along with sleep-related questions. The format for the sleep-related questions is a mixture of short answer write-ins, yes/no checkboxes, and multiple choice options. The sleep items address symptoms that are associated with insomnia, obstructive sleep apnea, restless legs syndrome, circadian rhythm disorders (delayed sleep phase), and parasomnias (nightmares, sleep walking, sleep talking). For the sleep items, the timeframe of interest is not specified, though for anxiety and depression the queries pertain to the last two weeks. The ASQ clusters items into components titled insomnia, mood, obstructive sleep apnea, delayed sleep phase disorder, and parasomnias. Sleep-related questions were based on the 2001 version of international classification of sleep disorders (ICSD) [29]; clustering of items was based on expert opinion. Reported values of sensitivity and specificity for insomnia, obstructive sleep apnea, delayed sleep phase disorder, and parasomnias were based on an expert opinion (an in-depth interview by a psychiatrist trained in sleep disorders).

Global sleep assessment questionnaire (GSAQ)

The GSAQ [24] comprises 11 items in grid format with responses on a single page. The top of the page includes queries regarding age, sex, height, weight, and employment status (not included in the 11-item count). Each of the 11 sleep items is constructed as a question next to a row of checkbox response options regarding symptom frequency over the last four weeks (never, sometimes, usually, and always). The 11 items cover mood, life activities and medical issues as they relate to sleep, along with symptoms associated with insomnia, obstructive sleep apnea, restless legs syndrome/periodic limb movement, and parasomnias. Note: the paper introducing the instrument suggested that the scale also covers circadian rhythm disorders. This is true to the extent that the header contains information about shift work and jet lag. This said, the more standard intrinsic sleep disorders of delayed and advanced sleep phase are not captured. Item construction was based on expert opinion. An initial list of items was screened using logistic regression and then decreased to the final item set based on confirmatory multivariable regression. Sensitivity and specificity values for insomnia, obstructive sleep apnea, restless legs syndrome/periodic limb movement, and parasomnias were based on clinical assessments (the judgment of expert sleep clinicians). AUC receiver operator characteristics were used to determine scale cut-points.

Holland sleep disorders questionnaire (HSDQ)

The HSDQ [25] comprises 32 items arrayed over four pages. The top half of the first page provides instructions for filling out

Fig. 1. Sleep disorders screening questionnaire literature search data extraction – candidates for additional appraisal of practice related features and psychometric qualities. Articles were counted according to the number of tools on which they reported (e.g., a reference that reported on three screening tools was tabulated as three "articles". Grey literature sources were identified via Google Scholar, Agency for Healthcare Research and Quality [18], Centers for Disease Control and Prevention [19] & American Academy of Sleep Medicine [20].

the form and lists the meaning of the response choice numbers. Each item is a short statement next to a row of five numbers; subjects circle the number representing the extent to which statements apply to them. The five numbers correspond to 1 = not at all applicable, 2 = usually not applicable, 3 = applicable at times, 4 = usually applicable, and 5 = applicable. The timeframe of interest was specified as being within the last three months. Each item was constructed based on expert opinion and was framed according to ICSD-2 [30] diagnostic criteria for insomnia, sleep disordered breathing, restless legs syndrome/periodic leg movement disorder, circadian rhythm sleep disorders, and parasomnias. While questions related to excessive sleepiness and sleep attacks are included, there are no items related to the remaining defining symptoms of narcolepsy (cataplexy, sleep paralysis, and hypnagogic or hypnopompic hallucinations). HSDQ development began with 40 items which were then pared down to 32 items using factor analysis and reliability data. Confirmatory factor analysis was used to check item grouping. Receiver operator characteristics analyses were used to determine cut-point scores for each of the six subscales, with the comparators being: a) diagnosis provided by sleep center for sleep center patients, and b) absence of sleep disorders (for the control group) based on experts' knowledge of the subjects recruited as controls.

Table 1
Summary of comprehensiveness, brevity, and time period assessed of sleep disorders screening questionnaires.

Questionnaire	# of items: response choices	#Sleep disorders	Time period assessed
ASQ (2011) [23] Auckland sleep questionnaire	30: yes/no	4: I, O, P, C	Not specified
SSC (2008) [22] Sleep symptom checklist	21: 0–3[a]	3[f] I, O, R	Past month
SDQ (1994) [27] Sleep disorders questionnaire	175: 1–5[b]	3: O, N, R	Past six months
HSDQ (2013) [25], Holland sleep disorders questionnaire	32: 1–5[c]	5: I, O, R, P, C	Past three months
ISDI (2010) [26] Iowa sleep disturbances inventory	86: yes/no	3[g] I, R, P	Not specified
GSAQ (2002) [24], Global sleep assessment questionnaire	11: 0–4[d]	4: I, O, R, P	Last four weeks
Sleep-50 (2005) [28]	50: 1–4[e]	6: I, O, N, R, P, C	Last four weeks

Abbreviations: C − circadian rhythm sleep wake disorders; DSM-diagnostic and statistical manual of mental disorders; I − insomnia; ICSD − international classification of sleep disorders; N − narcolepsy; O − OSA; P − parasomnias; R − restless legs syndrome; SD − sleep disorder.

Notes:
[a] Severity 0–3.
[b] Never [strongly disagree], rarely [disagree], sometimes [not sure], usually [agree], always [agree strongly].
[c] Not at all applicable, usually not applicable, applicable at times, usually applicable, applicable.
[d] Never, sometimes, usually, always.
[e] Not at all, somewhat, rather much, very much.
[f] Three SDs covered but grouped into non-SD-specific factors.
[g] Three SDs covered but grouped into non-SD-specific factors.

Iowa sleep disturbances inventory (ISDI)

The ISDI [26] comprises 86 items arrayed over six pages. Each item is a short statement next to which subjects check true or false. The timeframe of interest (regarding the incidence of symptoms) is not specified. The 86 items were pared down from large item pool of more than 3000 items (crafted by sleep disorders specialists at the University of Pittsburgh). The items were crafted to assess sleep disorders as defined by the ISCD-2 [30] and the DSM-IV (diagnostic and statistical manual of mental disorders) [31]. An exploratory factor analysis was used to select the final item set and as a starting point for item clustering. Extensive statistical analyses (additional factor analyses and regressions) were performed with multiple subject samples to determine the ten factor structure (nightmares, insomnia, fatigue, fragmented sleep, non-restorative sleep, anxiety at night, light sleep, movement at night, sensations at night, excessive sleep, and irregular schedule) which was then clustered into two higher order constructs. Correlation coefficients were calculated for the ISDI symptom clusters and other established questionnaires (e.g., Epworth sleepiness scale and the PSQI) to assess the validity of the ISDI.

Sleep disorders questionnaire (SDQ)

The SDQ [27] comprises 175 items arrayed over 12 pages. The first 152 questions are short statements next to a row of numbers 1–5 (never [strongly disagree], rarely [disagree], sometimes [not sure], usually [agree], always [agree strongly]), while the remaining 23 items are five level quantitative multiple choice questions (i.e., regarding frequencies, time intervals, incident rates, etc.). The time frame for all questions was the "past six months". Items were extracted from an earlier sleep disorder screener (the SQAW) by the same author group and then reworded based on expert opinion. Item content covers sleep apnea, restless legs syndrome/periodic limb movement disorder, and narcolepsy, as well as a general category denoted as psychiatric sleep disorder. Item grouping was determined by multivariate analysis. Cut-point scores were determined by receiver operator characteristics (ROC) analyses, with sensitivity and specificity values reported. ROC results and reported values of sensitivity and specificity were based on presence of sleep disorders determined by polysomnography or multiple sleep latency tests and absence of sleep disorders (in healthy controls) determined by expert interviews.

Sleep-50

As the name implies, this instrument comprises 50 items. Each item is a short statement which the subject may endorse on a 4-point-scale (1 [not at all], 2 [somewhat], 3 [rather much], and 4 [very much]). The time frame for all questions was the "past four weeks". The statements and scales are aligned in columns (a grid format) with responses arrayed over two pages. Items were based on expert opinion to cover DSM-IV [31] criteria for sleep disorders. Additionally, the instrument includes statements regarding potential moderating factors (e.g., questions re: common sleep hygiene infractions, mood disturbance, sleep effort and preoccupation, etc.). The instrument assesses for insomnia, obstructive sleep apnea, narcolepsy, restless legs syndrome, circadian rhythm disorders, and parasomnias. Items were initially grouped according to their respective sleep disorders. Principal components analysis was then used to guide a slight readjustment in clustering. Cut-point scores for each of the subscales was selected to give optimal sensitivity and specificity based on sleep clinic diagnoses via polysomnography, detailed physical/psychiatric history, sleep diaries, and unstructured interviews.

Sleep symptom checklist (SSC)

The SSC [22] comprises 21 items in grid format arrayed over one page. Each item is a symptom term (e.g., insomnia) or phrase (e.g., waking often to urinate) which the subject may endorse on a 4-point-scale with respect to severity (0 = not at all, 1 = mildly, 2 = moderately, 3 = very). The time frame for all questions was the "past month". Items were constructed based on expert opinion informed by previously published sleep diaries and earlier studies of geriatric sleep issues to cover insomnia, restless legs syndrome and obstructive sleep apnea. Additional items were included to assess for the daytime consequences of sleep disturbance. Factor analysis was used to group items into four components: insomnia, daytime aspects (daytime consequences of sleep disturbance), sleep disorder (any form of sleep disturbance), and psychological maladjustment (e.g., depression and anxiety). Polysomnography was utilized to establish presence or absence of sleep disorders. Scores on the four component subscales were correlated with presence or absence of each of the sleep disorders as assessed with polysomnography.

Quality appraisal results

None of the identified questionnaires scored 14 out of 14 possible points. Table 2 provides the scores for each of the seven questionnaires for comprehensiveness, brevity, and for each of the four questions pertaining to psychometric soundness. Only the sleep-50 was found to assess for all six of the intrinsic sleep

Table 2
Overall quality appraisal of multi sleep disorders screening questionnaires

Feature	Questionnaire / 1st author						
	ASQ (2011) [24] Auckland Sleep Questionnaire	*SSC (2008)* [23] Sleep Symptom Checklist	*SDQ (1994)* [28] Sleep Disorders Questionnaire	*HSDQ (2013)* [26] Holland Sleep Disorders Questionnaire	*ISDI (2010)* [27] Iowa Sleep Disturbances Inventory	*GSAQ (2002)* [25] Global Sleep Assessment Questionnaire	*SLEEP-50 (2005)* [29]
	Arroll	*Bailes*	*Douglass*	*Kerkhof*	*Koffel*	*Roth*	*Spoormaker*
Comprehensiveness[a]	4	3	3	5	3	4	6
Brevity[b]	0	2	0	0	0	4	0
Were all of the **items standardized (for wording)**, i.e., adopted from previously established instruments or ratified by recognized expert consensus panels?[c]	0	0	0	0	0	0	0
Was an effort made to **streamline the instrument using a statistical technique** (such as, confirmatory factor analysis or a logistic regression, for example, were items with low weightings were eliminated)?[c]	0	0	1	1	1	1	1
Was the item cluster (subscale) for each diagnosis shown to be relatively **sensitive and specific for the target disorder**, i.e., did the authors provide true/false positive rates, sensitivity/specificity rates, and/or ROC/AUC analyses?[c]	1	0	1	1	0	1	1
Was the instrument **evaluated for sensitivity to change in patient status**, i.e., spontaneous recovery or treatment-related symptom improvement?[c]	0	0	0	0	0	0	0
Grand Total[d]	5	5	5	7	4	10	8

Notes:
[a] one point per intrinsic sleep disorder
[b] > 25 items − zero points, 20-25 items − two points, <20 items − four points
[c] zero points if no, one point if yes
[d] up to 14 points

disorders and only the SSC and GSAQ were deemed to be brief. Most of the questionnaires had solid foundations in two of the psychometric areas: a) streamlining the number of items and grouping them into clusters (or subscales) and b) assessment of subscale predictions against target disorders. None of the questionnaires utilized standard item queries or were evaluated for their sensitivity to change in patient status. The highest composite score was for the GSAQ (score = 10) followed by the sleep-50 (score = 8).

Discussion

Seven self-report questionnaires were identified that assess for three or more of the six core intrinsic sleep disorders. While none of the questionnaires scored well for both comprehensiveness and brevity, the GSAQ scored highest overall for the quality appraisal. The GSAQ [24] is therefore identified as the most suitable instrument for use as a screener for primary care. The relative merits of the GSAQ are explicated below.

Strengths and limitations of the GSAQ

The primary strengths of the GSAQ [24], in addition to efficiency and comprehensiveness, are its physical layout and the effort to embrace factors that may predispose, precipitate, or perpetuate sleep disturbance (i.e., pain, medications, mood disturbance, etc.). The physical layout of the GSAQ is in many ways ideal because it: 1) is presented on a single page; 2) has a header that contains pertinent information regarding patient characteristics (e.g., age, sex, height and weight, work status, and work shift); 3) lays out symptom complaints within a grid that allows for easy patient response and clinician interpretation (i.e., the more checkboxes on the right the more sleep disorders morbidity); and 4) arrays relevant symptomatology by row where each row corresponds to

either a specific sleep disorder or the report of impaired daytime function. Finally, the GSAQ [24] also covers daytime impairment and insufficient sleep disorder (items 3 and 4).

The primary limitations of the GSAQ [24], given the present review criteria, are that 1) it does not allow for the detection of circadian rhythm disorders (DSPS/ASPS) and narcolepsy and 2) its reliability (content and construct validity) is less well developed than most of the other instruments (See Table 2). With respect to circadian rhythm disorders, the omission of this class of sleep disorders with such questions as "I sleep better if I go to bed before 21:00 h and wake up before 4:30 h" may have been intentional as the items in the header of the instrument regarding work shift may have been thought to be adequate proxy questions. With respect to narcolepsy, the omission of this sleep disorder from a primary care screener may have been (and may still be) viewed as less problematic based on the fact that narcolepsy is a rare disorder (estimated prevalence of 1/3000 [32]). Thus, it may not be critical that this be assessed by primary care practitioners. This said, the ramifications of not screening for an infrequently occurring disorder needs to be considered. Narcolepsy, while admittedly rare, has a profound impact on the health and wellbeing of the affected individual and this may warrant the addition of two to four questions to screen for this disorder.

Extending beyond the present criteria, the GSAQ [24] has some significant untapped potential. First, the four item response selection is arrayed qualitatively (never, sometimes, usually, and always). While this has the advantage of having general meaning, the variance in interpretation between individuals may be problematic. For example, some may interpret "sometimes" as once every few months, others once every few weeks, and still others may interpret sometimes as once a week. This problem with interpretation might be avoided if the response selections were set to frequencies of clinical relevance (and this might also aid in

the determination of the clinical significance of the endorsement and provide a way to assess severity). For example, "never" could be changed to "rarely" and each qualitative term could be anchored (incidence per unit time) where: "Rarely" refers to less than once a month, "Sometimes" refers to once every two to four weeks, "Usually" refers to weekly, and "Always" refers to more than three times per week. Further, extending the scale to a five item response could provide for the differentiation between three to five times per week and more than five times per week, allowing for a reasonable way of scaling symptom severity. Second, the item responses are not assigned numeric values. If zero to five numbering was used for each of the four responses, this would allow for a cumulative morbidity index. Third, the instrument was designed to take into account "the past four weeks". While this provides a very reasonable time frame for the estimation of what constitutes persistent symptomatology, it does not allow for the distinction between acute and chronic forms of the various sleep disorders. While not easily done within the given format, an additional checkbox per symptom regarding chronicity would allow clinicians to make judgments about whether recent life/health changes precipitated the observed sleep disorders symptom and/or about whether or not to initiate, or to make a referral for, treatment (e.g., a checkbox for "This has been true for the last three or more months").

Given that the GSAQ [24] represents the best possible screener available, dissemination of this instrument to primary care practice will likely provide clinicians the ability to engage in early detection and the initiation of treatment to prevent the adverse effects of untreated sleep disorders. Widespread dissemination of the GSAQ [24], however, will need to be supplemented with the provision of treatment and/or specialist referral guidelines. Several resources for this purpose are readily available and should be provided as part of the dissemination process including (but not limited to) information regarding how to access on-line practice and treatment guidelines, sleep center locations and contact information, and provider directories [33–36] In addition, primary care clinicians should be provided access to patient education materials that can be provided prior to the initiation of treatment or referral [37–40].

Limitations of this review

The search strategy for identifying questionnaires may have limited scope of the results. Key items may have been missed by including only those articles written in English using three medical/nursing databases; grey literature items may also have been missed, e.g., by not purposefully searching websites others may consider promising. Conclusions regarding psychometrics may be limited by having only examined introductory articles.

The brevity (number of items) criterion used for assessing instruments may not be reasonable. Although based on available estimates of time to complete questionnaires and typical times spent waiting prior to primary care visits, the conversion of that information into a limit of 25 items was, in the end, arbitrary and therefore open to debate. A better method of judging brevity of an instrument might have been to measure completion times by representative patients in typical primary care settings. Making such measurements was, however, considered outside the scope of this review. It may be beneficial for future instrument developers to report completion time data for the benefit of practitioners and as a benchmark for other developers.

The weighting of the assessment criteria for appraising the questionnaires may also seem arbitrary. Each of the three categories, i.e., comprehensiveness (six points), brevity (four points), and psychometric soundness (four points), contributed about equally to the overall appraisal of the instruments evaluated here,

however with 10 points allocated to comprehensiveness and brevity (vs. four points for psychometric soundness), the overall evaluation is biased towards the first two factors. It could be argued, however, that the weighting could have been operationalized in a different manner, e.g., psychometric soundness (as a guarantee of the accuracy of the instruments) could and should have been substantially expanded so that this factor would represent the primary consideration regarding instrument quality. We elected to adopt the former strategy (emphasize comprehensiveness and brevity) given that 1) this is what will allow for feasibility of use in primary care, and 2) that the psychometric soundness of a screener seems less important given that further evaluation is likely following screening.

Conclusion

While the GSAQ [24] represents the best available screening tool for primary care practice, further development is needed. Development should take into account the need for such instruments to: efficiently assess for the full complement of intrinsic sleep disorders (including insufficient sleep disorder); acquire the basic demographic information that is relevant for assessment (age, sex, BMI, presence/absence of bed partner, work status and shift); provide for an assessment of the daytime consequences of disturbed sleep (including but not limited to daytime sleepiness); include scales that are anchored to clinically relevant frequencies (e.g., days per week); account for the temporal course of the presenting symptoms (i.e., at least the distinction between acute and chronic forms of the various sleep disorders); and provide for the quantification of cumulative sleep disorders morbidity.

Practice points

This review has identified a questionnaire (the GSAQ) which screens for multiple sleep disorders that:

- encompasses most of the core intrinsic sleep disorders;
- requires minimal time for patients to complete;
- provides evidence of psychometric soundness;

This instrument assesses insomnia, obstructive sleep apnea, restless legs syndrome and parasomnias in eleven questions.

Research agenda

Further development of a questionnaire to screen for multiple sleep disorders will incorporate:

- query timeframes that mimic diagnostic criteria;
- response options anchored to clinically relevant frequencies;
- format and layout that are easily completed by patients and interpreted by providers;
- basic demographic information (age, sex, BMI, presence/absence of bed partner, work status and shift) and limited data regarding sleep sufficiency and the daytime consequences of sleep disturbance.

Conflicts of interest

Acknowledgments

References

*[1] World Health Organization. Global surveillance, prevention and control of chronic respiratory diseases: a comprehensive approach. 2007.

*[2] Institute of Medicine. Sleep disorders and sleep deprivation: an unmet public health problem. Washington, DC: National Academies Press; 2006. Available from: http://www.ncbi.nlm.nih.gov/books/NBK19958/.

[3] American Academy of Sleep Medicine. National healthy sleep awareness project. 2014.

*[4] Buysse DJ, Yu L, Moul DE, Germain A, Stover A, Dodds NE, et al. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. Sleep 2010;33:781–92.

[5] Rush AJ, Giles DE, Schlesser MA, Fulton CL, Weissenburger J, Burns C. The inventory for depressive symptomatology (IDS): preliminary findings. Psychiatry Res 1986;18:65–87.

[6] Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The inventory of depressive symptomatology (IDS): psychometric properties. Psychol Med 1996;26:477–86.

[7] Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck depression inventory: twenty-five years of evaluation. Clin Psychol Rev 1988;8:77–100.

[8] Beck AT, Ward C, Mendelson M. Beck depression inventory (BDI). Arch Gen Psychiatry 1961;4:561–71.

[9] Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278–96.

[10] Yesavage JA, Brink T, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1983;17:37–49.

[11] Kroenke K, Spitzer RL, Williams JB. The PHQ9: validity of a brief depression severity measure. J General Intern Med 2001;16:606–13.

[12] Morin CM, Belleville G, Bélanger L, Ivers H. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 2011;34:601.

[13] Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med 1999;131:485–91.

[14] Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. J Am Soc Anesthesiol 2008;108:812–21. –21.

[15] Validation of the international restless legs syndrome study group rating scale for restless legs syndrome. Sleep Med 2003;4:121–32.

[16] Johns MW. Reliability and factor analysis of the Epworth sleepiness scale. Sleep 1992;15:376–81.

[17] Buysse DJ, Reynolds 3rd CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213.

[18] Agency for Healthcare Research and Quality. National guideline Clearinghouse. 2014.

[19] Centers for Disease Control and Prevention. National health and nutrition examination Survey. Atlanta, GA: CDC; 2014.

[20] American Academy of Sleep Medicine. 2014.

[21] Anderson RT, Camacho FT, Balkrishnan R. Willing to wait? the influence of patient wait time on satisfaction with primary care. BMC Health Serv Res 2007;7:31.

*[22] Bailes S, Baltzan M, Rizzo D, Fichten CS, Amsel R, Libman E. A diagnostic symptom profile for sleep disorder in primary care patients. J Psychosom Res 2008;64:427–33.

*[23] Arroll B, Fernando A, Falloon K, Warman G, Goodyear-Smith F. Development, validation (diagnostic accuracy) and audit of the Auckland sleep questionnaire: a new tool for diagnosing causes of sleep disorders in primary care. J Prim Health Care 2011;3:107–13.

*[24] Roth T, Zammit G, Kushida C, Doghramji K, Mathias SD, Wong JM, et al. A new questionnaire to detect sleep disorders. Sleep Med 2002;3:99–108.

*[25] Kerkhof GA, Geuke MEH, Brouwer A, Rijsman RM, Schimsheimer RJ, Van Kasteel V. Holland sleep disorders questionnaire: a new sleep disorders questionnaire based on the international classification of sleep disorders-2. J Sleep Res 2013;22:104–7.

*[26] Koffel E, Watson D. Development and initial validation of the Iowa sleep disturbances inventory. Assessment 2010;17:423–39.

*[27] Douglass AB, Bornstein R, Nino-Murcia G, Keenan S, Miles L, Zarcone Jr VP, et al. The sleep disorders questionnaire. I: creation and multivariate structure of SDQ. Sleep 1994;17:160–7.

*[28] Spoormaker VI, Verbeek I, Van den Bout J, Klip EC. Initial validation of the SLEEP-50 questionnaire. Behav Sleep Med 2005;3:227–46.

[29] The International classification of sleep disorders. revised: diagnostic and coding manual. Chicago, IL: American Academy of Sleep Medicine; 2001.

[30] American Academy of Sleep Medicine. International classification of sleep disorders: diagnostic & coding manual. 2nd ed. Darien, IL: American Academy of Sleep Medicine; 2005.

[31] American Psychiatric Association. DSM-IV: diagnostic and statistical manual of mental disorders. Wahington, DC: American Psychiatric Association; 1994.

[32] National Institute of Neurological Disorders and Stroke. Narcolepsy fact sheet. 2015.

[33] American Academy of Sleep Medicine. AASM resource library. 2014.

[34] American Academy of Sleep Medicine. Practice guidelines. 2014.

[35] American Academy of Sleep Medicine. Find a sleep center (search box). 2014.

[36] Society of Behavioral Sleep Medicine. Find a BSM provider. n.d.

[37] National Heart Lung and Blood Institue. Patient & public information. n.d.

[38] National Sleep Foundation. Sleep disorders problems. 2014.

[39] National sleep foundation. Sleep health. n.d.

[40] National Institute of Neurological Disorders and Stroke. Brain basics: understanding sleep. 2014.

---

* The most important references are denoted by an asterisk.