pii: jc-00307-15 http://dx.doi.org/10.5664/jcsm.5676

JCSM
Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

A Three-Item Instrument for Measuring Daytime Sleepiness: The Observation and Interview Based Diurnal Sleepiness Inventory (ODSI)

Fannie Onen, MD, PhD[1,2,3]; Christophe Lalanne, PhD[2]; Victoria M. Pak, MS, PhD[3]; Nalaka Gooneratne, MD, MSc[3,4]; Bruno Falissard, MD, PhD[2]; Saban-Hakki Onen, MD, PhD[4,5]

[1]CHU Bichat Claude Bernard, APHP, Service de Gériatrie, Paris, France; [2]Université Paris Descartes, INSERM, Paris, France; [3]Center for Sleep and Circadian Neurobiology and Division of Sleep Medicine, School of Medicine, University of Pennsylvania, Philadelphia, PA; [4]Division of Geriatric Medicine, School of Medicine, University of Pennsylvania, Philadelphia, PA; [5]Hôpital Eduard Herriot, Centre Gériatrique de Médecine du Sommeil, Lyon, France

Study Objectives: We aimed to develop a new three-item assessment tool for daytime sleepiness in older adults, the Observation and interview-based Diurnal Sleepiness Inventory (ODSI) and determine its validity, internal consistency, test-retest reliability, and optimal cutoff score.

Methods: A total of 133 elderly subjects including 73 patients with obstructive sleep apnea (OSA) (mean age, 79 y) and 60 controls (mean age, 80 y) were consecutively enrolled and answered all questionnaires. The ODSI questionnaire was validated using the Epworth Sleepiness Scale considered as a gold standard. Reliability, validity, and cut-points were tested.

Results: The ODSI has acceptable validity, internal consistency, and test-retest reliability properties. The ODSI has internal consistency and a reliability coefficient (Pearson rho) of 0.70 for its three items, which suggests strong reliability. The estimated sensitivity and specificity were 0.842 with 95% confidence interval [0.624; 0.945] and 0.851 [0.761; 0.911], respectively. The consistency of summated scale scores during test and retest sessions was high (r = 0.970, 95% bootstrap confidence interval [0.898; 0.991]). Receiver operating characteristic analysis suggests that a cut-point of 6 is effective for identifying older adults with excessive levels of daytime sleepiness.

Conclusions: The ODSI is a brief, valid, easy-to-administer three-item assessment that can screen for excessive daytime sleepiness among elderly patients with OSA.

Keywords: excessive daytime sleepiness, obstructive sleep apnea, ODSI, older adult

Citation: Onen F, Lalanne C, Pak VM, Gooneratne N, Falissard B, Onen SH. A three-item instrument for measuring daytime sleepiness: the Observation and Interview Based Diurnal Sleepiness Inventory (ODSI). *J Clin Sleep Med* 2016;12(4):505–512.

INTRODUCTION

Aging modifies sleep-wake rhythms and can result in a diminution of nocturnal sleep associated with frequent naps during the day.[1] In the elderly, it may be difficult to distinguish physiological early afternoon naps and excessive daytime sleepiness. Excessive daytime sleepiness (EDS) is an extremely common problem, especially in older adults, that is associated with reduced alertness, impaired cognitive function, falls and an increased cardiovascular morbidity and mortality rate.[2–6] Even though EDS is a common clinical symptom in several medical conditions, there are currently no validated assessment tools for older adult populations.

Remarkably, most individuals fail to even report the symptom of daytime sleepiness to physicians, either because of a lack of recognition that the symptom is medically significant or a concern that it may be interpreted pejoratively as a sign of laziness or even senility. Even when the patient does register the complaint, however, physicians are unlikely to obtain a sleep history and are even less likely to attempt to assess the problem.[7]

Several methods currently exist to assess habitual sleepiness. The Epworth Sleepiness Scale (ESS) is the most commonly used subjective scale in clinical practice even in geriatrics. It asks the patient to estimate retrospectively how likely he/she is to doze off or fall asleep in eight different everyday situations for an active and autonomous adult.[8] In cases where patients have cognitive impairment, the clinical reality is that the ESS often has to be administered by the interviewer. Although eight items are brief for most assessments, they can be time-consuming for an interviewer to administer within the time constraints of a typical office visit encounter. Furthermore, there is often a discrepancy between patient report on the ESS and proxy report, such as a spouse or caregiver.[9,10] There is thus a need for a simple, standardized clinical test for measuring an

BRIEF SUMMARY

Current Knowledge/Study Rationale: Excessive daytime sleepiness is a common problem in older adults yet there are no currently validated scales for measurement of sleepiness in the elderly. The ODSI is a brief, valid, easy-to-administer three-item interviewer-guided tool that can screen for excessive daytime sleepiness. The ODSI score appears to have potential for clinical utility; it is usually 6 or above in patients with OSA syndrome that causes hypersomnolence but averages 3 in healthy controls.

Study Impact: As the clinical diagnosis of OSA is increasingly being made by primary care physicians with limited time to assess behavioral morbidity, the ODSI will be an important method for measuring sleepiness in clinical care.

older patient's daytime sleepiness duration and general level of sleepiness in daily activities that can also include information obtained from a proxy.

We have developed a tool for the rapid assessment of EDS in older adults that can incorporate proxy input. We have termed this instrument the Observation and Interview Based Diurnal Sleepiness Inventory (ODSI). The instrument is currently available in two languages (French and English) and has been used successfully within our centers (Lyon, Paris, and Philadelphia) to evaluate the sleepiness in older adults in routine clinical practice and ongoing research.

The current report describes the development and validation of the ODSI, including internal consistency, cut-point, test-retest reliability, and testing the convergent validity of the ODSI in relation to another widely used measure, the ESS.

METHODS

The study was carried out according to the Declaration of Helsinki and approved by the Local Ethics Committee. Each subject included in the study gave informed consent.

Study Protocol

A total of 133 older subjects including 73 patients with OSA and 60 controls were consecutively enrolled. Each subject underwent cognitive (Mini Mental State Examination) and functional (instrumental activities of daily living) assessments as well as height and weight measurements (used to calculate the body mass index [BMI]). The following comorbid conditions were checked for all subjects (medical history by interview and chart abstraction): hypertension, stroke and transient ischemic attack, ischemic heart disease, atrial fibrillation, and diabetes mellitus.

Subjects

Control subjects

The control subjects were recruited from senior citizen centers and from among people visiting a relative (wife, husband, sister, brother) in our geriatric units in Paris and Lyon.

The control subjects were eligible if they met the following inclusion criteria: (1) 65 y old or older; (2) ability to speak French fluently; (3) Caucasian race (the study sample was limited to Caucasian patients as the prevalence of non-Caucasian patients in our clinic was too low to allow for meaningful subgroup analyses by race/ethnicity); (4) living at home; (5) ability to drive or take public transportation without assistance. Inclusion criteria with answers in the negative to all of the following questions: (1) Do you often feel sleepy during daytime? (2) Have you recently had any trouble sleeping at night (insomnia, hypersomnia, loud snoring)? (3) Have you recently used medicine to help you sleep? (4) Have you recently used medicine to help you stay awake? (5) Have you ever been diagnosed with and/or treated for sleep apnea? (6) Has anyone ever witnessed you stop breathing while sleeping? Additional exclusion criteria were: (7) obesity defined as a BMI > 30 kg/m^2; (8) dementia and/or major psychiatric disorder; (9) psychotropic medicine use.

Obstructive sleep apnea syndrome patient population

Clinical data were obtained from patients with obstructive sleep apnea syndrome (OSAS) admitted consecutively in our geriatric sleep centers (Paris and Lyon). Patients were referred due to clinical suspicion of sleep apnea either by family practitioners (patients living at home) or by physicians from geriatric hospitals with acute care, rehabilitation, long-term care facilities, and outpatient clinics.

Clinical suspicion of OSAS was based on complaints of (1) loud snoring or witnessed apneas, (2) excessive daytime sleepiness, or (3) overweight/obesity. At least one of these symptoms was reported by the patient, relatives, or caregivers during a structured interview: (1) Do you snore loudly? (2) Has anyone ever witnessed you stop breathing while sleeping? (3) Do you often feel sleepy during daytime? Overnight polysomnography was arranged for all patients within 1 mo of completing the questionnaires.

The first 73 patients who met the following inclusion criteria were enrolled: (1) age 65 y and older (2) being Caucasian (the study sample was limited to Caucasian patients as the prevalence of non-Caucasian patients in our clinic was too low to allow for meaningful subgroup analyses by race/ethnicity) (3) apnea-hypopnea index (AHI) ≥ 15 and diagnosis of OSAS according to the International Classification of Sleep Disorders, Second Edition.[11]

The exclusion criteria were heart failure, nocturnal oxygen supplementation, dementia, major psychiatric disorder, being too sick to be evaluated, or another condition preventing the use of polysomnography. Patients were also not included in this study if they had previously undergone a sleep study or received care of any type for proven or suspected sleep apnea, if their final diagnosis was not OSAS, or they were unable to give informed consent.

Study Assessments

ODSI: Development and description of the ODSI

The ODSI is a brief, three-item assessment tool. The ODSI items emerged from a review of the literature and expert consensus in 2008.[12–18] The experts were two geriatricians (FO, SHO) involved in sleep medicine and having experiences in relation with causes, consequences, and severity of daytime sleepiness in older attending outpatient clinics, daily care clinics, and acute as well as long-term care units. A pragmatic consensus of the questions and standardized protocol was reached after in-depth discussion with caregivers and nurses in Paris and Lyon. Three items were proposed after expert consensus in order to assess sleep propensity and daytime sleepiness duration in an older subject's everyday life without significant cultural, scholarly, or professional references.

The ODSI is designed to be administered as an interview to patient and/or proxy. One appropriate way to check an elderly patient's sleepiness is to ask not only the patient but also interview relatives and caregivers who usually observe him/her in the day. These relatives (proxy) and caregivers often describe how their elder charge may doze off inadvertently and report how long they nap.

Each item is rated on a seven-point Likert scale and the total score ranges from 0 to 24. The first item examined sleepiness

during basic activities of daily living. The second item related to falling as asleep during periods of inactivity. The third item asked about hours of daytime sleep.

The first item, scored from 0 to 12, helps assess sleepiness or falling asleep in active situations. Falling asleep or dozing inadvertently during activities requiring high levels of stimulation exposes patients to danger with increased risk of falls and fractures and motor vehicle accidents.[3,19] Nodding off while driving a vehicle even rarely or mildly is more dangerous than falling asleep very frequently while watching TV (the second question) and taking voluntary naps and/or waking up late with a daytime sleepiness reaching 6 h (the third item). For both the second and third items, six is the maximum score. Because falling asleep during active situations (the first item) is likely more dangerous, experts proposed to add 6 to the first item while keeping a "0" unchanged. Thus, scoring for the first item would be 0, 7 (1+6), 8 (2+6), etc. and scores would range from 0 to12. In other words, with this scoring system, any positive answer to the first item is clinically more relevant than any other positive answer to the second and third items.

The second item scored from 0 to 6 helps assess sleepiness or falling asleep during passive situations, such as while reading or watching TV. Usually, these are not dangerous conditions and do not have a major effect on the patient's security. The third item helps estimate the average total duration of sleep during the day including sleepiness, falling asleep, and naps.

Furthermore, a scoring system with intermediate values allows for a quick consensus when patients and care givers are interviewed simultaneously. A similar scoring system have been used in Montgomery-Asberg Depression Rating Scale depression questionnaire.[12] The total score ranges from 0 (no somnolence) to 24 (excessive somnolence).

The questions were formulated in English and in French by the bilingual French geriatricians (FO, SHO). A bilingual American sleep medicine specialist (NG) examined and approved after minor modifications the conceptual structure of the English and French text. The data presented in this manuscript relates to the French version.

Description of the ESS

The ESS is an eight-item self-administered questionnaire asking the subject how likely he/she is to doze off or fall asleep in different situations of everyday life including (1) sitting and reading, (2) watching TV, (3) sitting inactive in a public place (theater, meeting), (4) as a passenger in a car for 1 h or longer, (5) when lying down to rest in the afternoon when circumstances permit, (6) when sitting and talking to someone, (7) when sitting quietly after lunch without alcohol, and (8) in a car while stopped for a few minutes in traffic. For each item, the subject may report the chance of dozing as never (score = 0), slight (score = 1), moderate (score = 2) or high (score = 3). Total scores range from 0 to 24, and an ESS score greater than 10 is considered as suggestive of pathologic sleepiness.[14] Subjects were asked to complete the ESS to provide a self-report of their sleepiness level.

The ESS is the most commonly used tool to assess EDS in older adults in both clinical practice and major cohort studies.[20–23] However, in older adults, unanswered items might compromise the relevance of the ESS score.[10]

ODSI and ESS administration

All control subjects and OSA patients were assessed with the ODSI and the ESS questionnaires. In this validation study, the ODSI was administered by an interviewer, who was also a rater (rater A). The rater A asked the subject the questions and then rated him/her based on the answers given by him/her and any pertinent additional information supplied by a relative or proxy. Another rater (rater B), present at the interview, with equal access to the same information about the subject only listened during the interview and then rated the subject. Neither rater had any information about how the other rater rated the subject. The investigators were randomly selected to be either rater A or rater B. The ODSI interviews were performed in the presence of at least one relative or proxy. These proxies were encouraged to comment on the subjects' answers with prompts such as "Do you agree with him/her?", or "What do you think about this answer?".

In order to evaluate test-retest reliability in a subsample of subjects, the ODSI was administered twice by the same interviewer 2 w later. This was a convenience sub- sample composed of subjects still available 2 w after the first assessment.

Polysomnography

The nocturnal polysomnogram was performed in our sleep centers. Using standard techniques,[24,25] a computer data acquisition and analysis system recorded the following signals: electroencephalogram (C3A2, C4A1, O1A2, and O2A1), bilateral electrooculogram, electrocardiogram, submental and bilateral anterior tibialis electromylogram, thoracic and abdominal excursion by piezocrystals, oral and nasal airflow by thermistor and breath sounds, body position, and oxygen saturation by pulse oximeter.

Statistical Methods

First, the questionnaire's reliability was determined by assessing the inner consistency and concordance using interitem correlations and weighted kappa statistics. The validity and quality was assessed by comparing answers to the ESS score. Next, a receiver operating characteristic analysis was used to determine the ODSI cut-off based on the ESS scale as a gold standard. All statistical tests were two-tailed and conducted with a type 1 error set at $p < 0.05$. Descriptive statistics were presented for demographic and clinical characteristics and the polysomnographic respiratory parameters.

Item analysis was done as follows: First, item scores were summarized using classic measures of central tendency (mean and median) and dispersion (standard deviation [SD] and interquartile range), controlling for floor and ceiling effects. Second, inter-item correlations (Pearson and Spearman coefficient) and correlations between item scores and summated scale score (including or excluding one of the three items each time) were examined to verify the consistency of the rating scale. The use of rest scores (total scores computed without including item under consideration) usually provides a more objective assessment of item-scale correlation, especially in the case of a short scale with unequal item weights. Third, the distribution of ODSI individual total scores was compared between patients and control, and with that of the ESS scores.

Table 1—Demographic and biomedical characteristics of participants.

	n	Control (n = 60)	OSA (n = 73)	Combined (n = 133)	Unadjusted p value
Age, y	133	80 (7) [75–85]	79 (7) [74–85]	80 (7) [74–85]	p = 0.7
Sex, female	133	73% (44)	41% (30)	56% (74)	p < 0.001
BMI, kg/m^2	133	26 (4) [23–29]	30 (7) [25–33]	28 (6) [24–31]	p < 0.001
Epworth Sleepiness Scale [a]	131	5 (3) [3–7]	10 (5) [6–13]	8 (5) [4–11]	p < 0.001
MMSE	133	28 (1) [27–29]	25 (4) [24–28]	27 (3) [26–29]	p < 0.001
IADL	132	7.8 (0.7) [8–8]	6 (2.4) [4–8]	6.8 (2) [6–8]	p < 0.001
Stroke	133	0% (0)	14% (10)	8% (10)	p = 0.003
Ischemic heart disease	133	7% (4)	23% (17)	16% (21)	p = 0.009
Atrial fibrillation	133	12% (7)	22% (16)	17% (23)	p = 0.1
Diabetes mellitus	133	10% (6)	22% (16)	17% (22)	p = 0.07
Hypertension	133	32% (19)	74% (54)	55% (73)	p < 0.001
Sleepiness complaint	73	–	74% (54)	–	–
Apnea-hypopnea index	73	–	45 (24) [27–59]	–	–
Average SpO_2	73	–	92% (3) [91–94]	–	–
Minimum SpO_2	73	–	75% (12) [61–85]	–	–
SpO_2 < 90%	73	–	16% (23) [2–17]	–	–
Sleep efficiency	73	–	66% (16) [54–78]	–	–
Arousal index	71	–	18 (13) [9–22]	–	–

n is the number of complete observations for each variable. Numerical variables are summarized as mean (standard deviation) [interquartile range]. Categorical variables are summarized as frequencies (counts). Two-sample comparisons were based on Wilcoxon and chi-square test for numerical and categorical variables, respectively. [a] For the Epworth Sleepiness Scale, only patients having completed all eight items were considered. BMI, body mass index; IADL, instrumental activities of daily living; MMSE, Mini Mental State Examination; OSA, obstructive sleep apnea syndrome defined with apnea-hypopnea index ≥ 15/h.

Fourth, a weighted kappa measure was used to assess test-retest reliability on a subset of the sample (22 control and 35 OSA) that were subjected to dual cotation; total scale scores were compared using "agreement" intraclass correlation computed from an analysis of variance table. Finally, the temporal stability of ODSI scores was assessed using "consistency" intraclass correlation computed from an analysis of variance on n = 19 available cases.

Concurrent validity was assessed by comparing ODSI and ESS scores of OSA patients free of cognitive impairment and with complete ESS responses to controls using linear and rank-based correlation coefficients. To verify known-group validity, average scores of OSA patients with and without daytime sleepiness complaints, and those of controls, were compared using Wilcoxon tests.

Receiver operator characteristic analysis was used to determine the ODSI cutoff based on the ESS, which served as a gold standard or reference test for classifying participants as cases (positives) or noncases (negatives) with respect to drowsiness. Participants were considered positive if they had a score of 11 or more on the ESS, and negative otherwise. Sensitivity (Se) and specificity (Sp) values, as well as the Youden index (Se+Sp-1), were computed for each ODSI cutoff score (0 to 24 points), and positive/negative predictive values, the score that offered the best compromise between these two equally weighted measures of diagnostic accuracy was retained. Positive and negative predictive values associated to the cutoff values were also reported with 95% asymptotic CI.

RESULTS

Subjects

A total of 133 older participants (60 controls and 73 patients with OSA) completed the ODSI questionnaire. The demographic and clinical characteristics of these participants are listed in **Table 1**.

The two groups were of similar age on average, but differed on all other characteristics, including sex (n = 19) and BMI (p < 0.001 in both cases). Seventy-four percent of patients (n = 54) with OSAS complained of excessive daytime sleepiness. In contrast, 26% of patients with OSAS and all control subjects (n = 60) did not complain of excessive daytime sleepiness.

Finally, 63% of patients with OSAS (n = 46) were able to answer all self-administered Epworth items. Thus, the ESS score analysis is based on 106 subjects (46 patients with OSAS and 60 controls).

The mean ODSI scores were highly significantly different (p < 0.001) between healthy nonselected older controls (2.6 ± 1.5) and older patients with OSAS (8.4 ± 5.2).

Internal Consistency

Item statistics

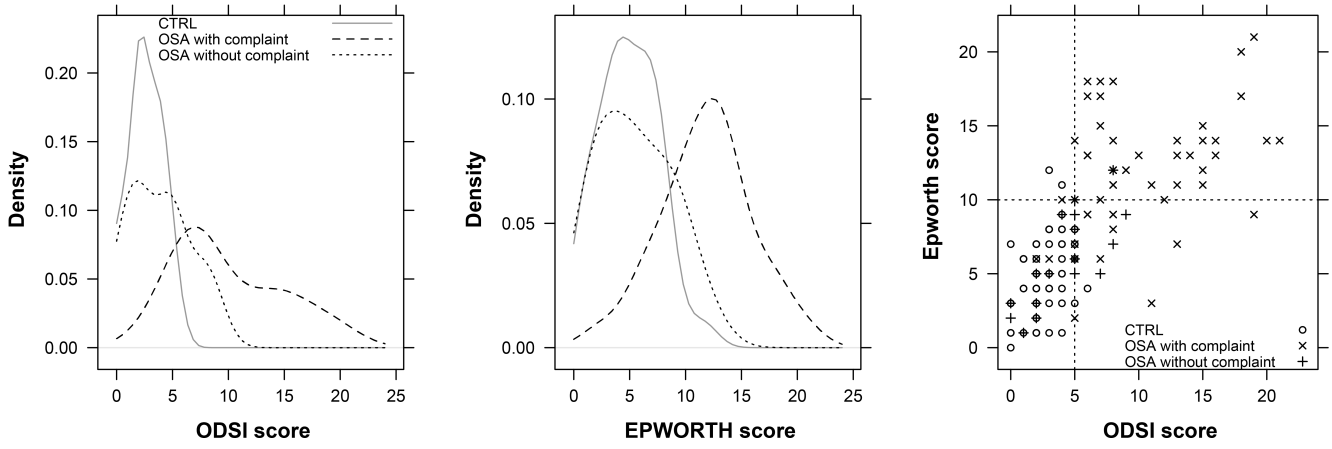
Item scores are summarized in **Table 2**. Floor effects (lowest response category chosen by a high number of respondents) were more apparent for the first item, with 84.2% of the participants having a score of 0. Every response category was observed for item 2, but responses were never above the fourth

Table 2—Descriptive statistics for items scores.

Item	Range	Mean (SD)	Median (IQR)	% (n) floor	Total score	Rest score	Kappa	Item 1	Item 2
Item 1	0–8	0.75 (1.95)	0 (0–0)	84.2 (112)	0.778	0.415	0.93	–	–
Item 2	0–6	3.39 (1.75)	4 (2–5)	8.3 (11)	0.812	0.534	0.984	0.348 (0.371)	–
Item 3	0–6	1.10 (1.28)	1 (0–2)	41.4 (55)	0.795	0.614	0.982	0.408 (0.351)	0.612 (0.692)

Total score: correlation of item scores with total score, Rest score: correlation of item scores with total score computed without the item. Kappa refers to the weighted measure of agreement between any two set of ordinal measures. Interitem correlations are given as Pearson r (Spearman rho). IQR, interquartile range; SD, standard deviation.

Figure 1—Distribution of Observation and Interview Based Diurnal Sleepiness Inventory (ODSI) and Epworth Sleepiness Scale scores.



The patients with obstructive sleep apnea (OSA) reporting sleepiness complaints are separated from those not reporting any complaint and from controls (CTRL). The joint distribution of Epworth Sleepiness Scale and ODSI scores is shown in the right panel for participants having completed all Epworth items.

response option ("moderately") for item 3, and for item 1 only scores of 0 (n = 95), or 7 to 10 (n = 11) were observed.

Item-scale correlation

Correlations of each item score with the full scale score were above 0.700 (**Table 2**), whereas correlations with test scores remained higher than 0.500 for all but one item (item 1). Items 2 ("falling asleep during periods of inactivity…") and 3 ("daily sleeping hours") were more correlated (r = 0.612) each other, compared to other item pairs.

Total score

The distribution of total scores for controls and patients is shown in **Figure 1** as non-parametric density curves, which provide a smooth approximation of histograms of counts. Overall, mean score was estimated at 5.8 (4.9), with a significant average difference of 5.8 points between controls (2.6 ± 1.5 points) and patients (8.4 ± 5.2 points; Wilcoxon test, $p < 0.001$).

Concurrent (criterion) validity

For comparison with the ESS, only patients with OSAS having completed all ESS items were considered (n = 46); hence, a sample of 106 complete cases (60 controls and 46 OSA patients).

Linear and monotonic correlation between ODSI and ESS scores were in the acceptable range (Pearson r = 0.697 with 95% bootstrap CI [0.576; 0.781], Spearman rho = 0.728 [0.607; 0.801]).

A total of 19 participants (17.9%), mainly from the OSA group (n = 17 out of 46) scored above 10 on the ESS. Participants with scores above the ESS clinical cutoff (10 points) had an average score of 9.6 points (SD, 4.4) on the ODSI questionnaire, compared to 3.5 points (SD, 2.7) for participants whose score was ≤ 10 points on the ESS. For participants in the expected "normal" range on the ESS, there was a significant difference between mean scores of controls and patients (Wilcoxon test, p = 0.007).

Known-group validity

Among the patients with OSA, those reporting sleepiness complaints generally have a higher ODSI score (10.0 ± 4.9) compared to those who do not (3.9 ± 2.8), and this difference proved to be significant (Wilcoxon test, $p < 0.001$). In comparison, controls scored on average 2.6 ± 1.5 points on the ODSI scale, which is not regarded as significantly different from patients with OSA without complaints (p = 0.091).

Considering the number of respiratory events per hour of sleep (as measured by AHI) as a marker of the severity of the OSA, we found a weak correlation with ODSI total score (Spearman rho = 0.193, p = 0.103), yet it was slightly higher for patients without sleepiness complaints (rho = 0.262) compared to patients with complaints (rho = 0.132). Correlations were also weak at the item level (item 1, 0.190; item 2, −0.003; item 3, 0.144). Average AHI level for patients with OSA without

Table 3—Measures of diagnostic accuracy with 95% confidence interval based on the binomial distribution.

Score	TP / TN	Se (95% CI)	Sp (95% CI)	PPV (95% CI)	NPV (95% CI)	PLR (95% CI)	NLR (95% CI)
2	37 / 9	1.000 (0.906–1.000)	0.181 (0.116–0.271)	0.325 (0.389–1.000)	1.000 (0.816–1.000)	1.22 (1.11–1.34)	–
3	37 / 17	1.000 (0.906–1.000)	0.415 (0.321–0.516)	0.402 (0.389–1.000)	1.000 (0.910–1.000)	1.71 (1.44–2.03)	–
4	37 / 39	0.973 (0.862–0.995)	0.543 (0.442–0.640)	0.456 (0.388–0.995)	0.981 (0.899–0.997)	2.13 (1.69–2.69)	0.05 (0.01–0.35)
5	36 / 51	0.946 (0.823–0.985)	0.681 (0.581–0.766)	0.539 (0.386–0.985)	0.970 (0.896–0.992)	2.96 (2.17–4.03)	0.08 (0.02–0.31)
6	**35 / 64**	**0.919 (0.787–0.972)**	**0.840 (0.753–0.901)**	**0.694 (0.394–0.972)**	**0.963 (0.898–0.988)**	**5.76 (3.58–9.26)**	**0.10 (0.03–0.30)**
7	34 / 79	0.784 (0.628–0.886)	0.872 (0.790–0.925)	0.707 (0.375–0.886)	0.911 (0.834–0.954)	6.14 (3.53–10.67)	0.25 (0.13–0.46)
8	29 / 82	0.703 (0.542–0.825)	0.904 (0.828–0.949)	0.743 (0.369–0.825)	0.885 (0.806–0.935)	7.34 (3.84–14.04)	0.33 (0.20–0.54)
9	26 / 85	0.514 (0.359–0.666)	0.947 (0.882–0.977)	0.792 (0.349–0.666)	0.832 (0.750–0.891)	9.65 (3.99–23.35)	0.51 (0.37–0.71)
10	19 / 89	0.460 (0.310–0.616)	0.957 (0.896–0.983)	0.810 (0.341–0.616)	0.818 (0.736–0.879)	10.79 (4.02–28.99)	0.56 (0.41–0.76)

Cutoff value in bold. Higher values for all indicators are generally better. CI, confidence interval; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value; Se, sensitivity; Sp, specificity; TP, true positive; TN, true negative.

complaint was 39.7 ± 20.2, whereas it was 46.9 ± 24.6 for those reporting complaints, but the difference was not statistically significant (Wilcoxon test, p = 0.302).

Spearman correlations between sleep efficiency and ODSI scores and between arousal index and ODSI scores were weak (rho = 0.208, p = 0.078 and rho = 0.184, and p = 0.124, respectively). It should be noted that those two indicators were less correlated to the ESS scores (sleep efficiency, rho = 0.153, p = 0.202; arousal index, rho = −0.052, p = 0.670). Sleep efficiency percentage was on average close for patients with (63.7 ± 17.3%) and without (62.3 ± 17.9%) complaints and not significantly different (Wilcoxon test, p = 0.981). Arousal index was on average close for patients with (17.0 ± 10.6) and without (16.7 ± 10.9) complaints and not significantly different (Wilcoxon test, p = 0.922).

For patients with OSA, we again found a weak correlation between saturation of peripheral oxygen (SpO_2) < 90% and ODSI scores (Spearman rho 0.241, p = 0.040). Percentage time spent at SpO_2 < 90% was higher for patients with complaints (18.9 ± 25.1%) compared to patients without complaint (8.9 ± 15.1%). However, on average minimum SpO_2 and average SpO_2 were respectively close for patients with (72.3% and 80.7%, respectively) and without (91.6% and 93.2%, respectively) complaints.

Interrater agreement and test-retest reliability after 2 w

Dual cotation was available for n = 57 participants (22 control and 35 OSA). Results for individual items are presented in **Table 2**. Rank correlation between the two series of total scores was high (Spearman ρ = 0.972, 95% bootstrap CI [0.903; 0.995]) suggesting that overall there was a consistent assessment between the two raters. The "agreement" intraclass correlation computed from an analysis of variance table was good (r = 0.856, 95% bootstrap CI [0.796; 0.902]).

The consistency of summated scale scores during test and retest session was high (r = 0.970, 95% bootstrap CI [0.898; 0.991]).

Determination of the cutoff

Of the 106 complete cases, a total of 19 participants (17.9% sample prevalence), including 17 of 46 patients with OSAS and 2 controls (out of 60), were screened as positive using ESS clinical cutoff. Results from receiver operating characteristic analysis are indicated in **Table 3** for key values around the chosen cutoff, which corresponds to a score of 6 points on the ODSI questionnaire. The estimated sensitivity and specificity for that value were 0.842 with 95% CI [0.624; 0.945] and 0.851 [0.761; 0.911], respectively, which gave a Youden index of 0.693. The area under the curve was 0.902, which is generally suggestive of good diagnostic accuracy. Similar results were found after controlling for age and BMI (area under the curve = 0.926, Se = 0.947, Sp = 0.828).

The associated positive predictive value suggests that, when using ODSI as a screening or diagnostic criteria, 55.2% of the sample would be classified in the same way as when using the ESS.

At the cutoff value of 6 points on the ODSI questionnaire, 60.9% (n = 28) of the OSAS patients would be screened as positive, whereas only one of the 59 controls would be considered a case.

DISCUSSION

The ODSI is a brief, easy to administer, three-item assessment that can screen for excessive daytime sleepiness in patients with OSA as compared to controls. It has acceptable internal consistency (reliability), validity, and high test-retest reliability properties. In addition, the correlation with the longer ESS is also acceptable at Pearson r = 0.697. Receiver operating characteristic analysis suggests that a cutoff point of 6 is effective for identifying older adults with excessive levels of daytime sleepiness.

Our data obtained in healthy subjects provide normal reference values with an upper limit of the normal range of 3.10. The ODSI shows a good internal consistency. Using Pearson rho coefficient, correlations of each item score with the full scale score were above 0.70. Based on our data, the minimal difference in ODSI scores that can be reliably detected is 0.15.

Our results support construct, convergent, and concurrent validity. The observed differences in scores between patients and controls were significant. Specifically, patients with OSA who reported complaints of sleepiness scored substantially

lower than controls and the difference was significant. Because the item intercorrelations for all item pairings are very high, this provides evidence that all items are related to the same construct of excessive daytime sleepiness and demonstrates convergent validity. Concurrent validity was measured using the ESS as a "gold standard" and there was sufficient agreement demonstrated. The ODSI's high correlation with the ESS suggests it correlates well with a previously validated measure.

Internal consistency and test-retest reliability was acceptable. The ODSI has internal consistency and a reliability coefficient (Pearson rho) of 0.70 for its three items, which is a strong reliability. This scale consists of one item on falling asleep during periods of inactivity (passive) and one item on daily sleeping hours. These were more correlated with each other (r = 0.612) than other item pairs (falling asleep during active situations). This may be due to the fact that falling asleep during active situations and daily sleeping hours is not normal, whereas falling asleep during passive situations would be more likely. The consistency of scores over time assessed using intraclass correlations with test-retest data was very good (intraclass correlation coefficient = 0.856). The consistency of summated scale scores during test and retest session was high (r = 0.970, 95% bootstrap CI [0.898; 0.991]), suggesting strong reliability.

Currently, the ESS is one of the most widely used assessment tools in sleep medicine. It was initially developed in younger cohorts and effectively distinguished sleepy and non-sleepy groups.[8,14] The ODSI differs from ESS in at least two ways: (1) The ODSI is conducted by a health professional as an interview with patients and informants and the ESS is a self-administered questionnaire. (2) The ODSI is composed of three items that may be universally adaptable (different cultures, different ages, subjects living at home or at institution), whereas the eight items of ESS are mostly related to the active young adult's life. A high degree of correlation of ODSI scores and ESS results could be demonstrated.

The ODSI seems to be a good assessment instrument for excessive daytime sleepiness. First, among the patients with OSA, those reporting sleepiness complaints have a significantly higher ODSI score compared to those who do not. Second, the ODSI scores are correlated with apnea severity (AHI and percentage time at $SpO_2 < 90\%$) and disturbed sleep quality (sleep efficiency and arousal index) measured by polysomnography.

The ODSI score appears to have potential for clinical utility; it is usually 6 or above in patients with OSAS that causes hypersomnolence but averages 3 in healthy controls. However, the range of statistically abnormal scores extends well beyond the cutoff of 6 used to define excessive sleepiness with ODSI. Future research should explore developing gradations of sleepiness (mild, moderate, or severe) based on different ODSI threshold values and possible impact of comorbid diseases that may lead to sleepiness even in the setting of lower AHI values.

Limitations

Some limitations of the current study must be acknowledged. Because our sample was limited to Caucasian respondents, future research with additional racial and ethnic groups is necessary. Because patients were recruited from mainly geriatric sleep centers, future studies should investigate the reliability of the ODSI in larger populations with other sleep disorders and conditions in primary care in order to encourage the use of the ODSI in diverse healthcare settings. Another limitation is that the gold standard was based on the ESS; therefore, the sensitivity of the questionnaire could be affected. Nevertheless, the literature has demonstrated the ESS is highly specific for sleepiness.[8,14] Finally, because of the nature of the ODSI questionnaire, it is subject to surrogate interviewer bias in which the person interviewed (i.e., spouse/child) may not have accurate information about the patient's history.

CONCLUSIONS

The ODSI represents the first standardized method for measuring sleepiness among the elderly that can incorporate proxy input using a brief interviewer-guided tool. In this paper, we present data supporting its reliability and validity for assessing sleepiness in this population. Based on the current results, a score of 6 or above in patients on the ODSI would be an appropriate cutoff score to suspect clinically relevant sleepiness among elderly clinic populations. As the clinical diagnosis of OSA is increasingly being made by primary care physicians with limited time to assess behavioural morbidity, the ODSI will be an important screening tool for clinical care.

ABBREVIATIONS

AHI, apnea-hypopnea index
BMI, body mass index
CI, confidence interval
EDS, excessive daytime sleepiness
ESS, Epworth Sleepiness Scale
IADL, Instrumental Activities of Daily Living
MMSE, Mini Mental State Examination
NLR, negative likelihood ratio
NPV, negative predictive value
PLR, positive likelihood ratio
PPV, positive predictive value
ODSI, Observation and Interview Based Diurnal Sleepiness Inventory
OSA, obstructive sleep apnea
OSAS, obstructive sleep apnea syndrome
SD, standard deviation
Se, sensitivity
Sp, specificity
TP, true positive
TN, true negative

REFERENCES

1. Ohayon MM. Interactions between sleep normative data and sociocultural characteristics in the elderly. J Psychosom Res 2004;56:479–86.
2. Ohayon MM, Vecchierini MF. Daytime sleepiness and cognitive impairment in the elderly population. Arch Intern Med 2002;162:201–8.
3. Onen F, Higgins S, Onen SH. Falling-asleep-related injured falls in the elderly. J Am Med Dir Assoc 2009;10:207–10.

4. Empana JP, Dauvilliers Y, Dartigues JF, et al. Excessive daytime sleepiness is an independent risk indicator for cardiovascular mortality in community-dwelling elderly: the three city study. Stroke 2009;40:1219–24.
5. Gooneratne NS, Richards KC, Joffe M, et al. Sleep disordered breathing with excessive daytime sleepiness is a risk factor for mortality in older adults. Sleep 2011;34:435–42.
6. Gooneratne NS, Vitiello MV. Sleep in older adults: normative changes, sleep disorders, and treatment options. Clin Geriatr Med 2014;30:591–627.
7. Haponik EF. Sleep disturbances of older persons: physicians' attitudes. Sleep 1992;15:168–72.
8. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep 1992;15:376–81.
9. Kumru H, Santamaria J, Belcher R. Variability in the Epworth sleepiness scale score between the patient and the partner. Sleep Med 2004;5:369–71.
10. Onen F, Moreau T, Gooneratne NS, Petit C, Falissard B, Onen SH. Limits of the Epworth Sleepiness Scale in older adults. Sleep Breath 2013;17:343–50.
11. American Academy of Sleep Medicine. International classification of sleep disorders: diagnostic and coding manual, 2nd ed. Westchester, IL: American Academy of Sleep Medicine, 2005.
12. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382–9.
13. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index (PSQI): a new instrument for psychiatric research and practice. Psychiatry Res 1989;28:193–213.
14. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991;14:540–5.
15. Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. Sleep 1997;20:835–43.
16. Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep 2004;27:1255–73.
17. Pack AI, Dinges DF, Gehrman PR, Staley B, Pack FM, Maislin G. Risk factors for excessive sleepiness in older adults. Ann Neurology 2006;59:893–904.
18. Onen SH, Dubray C, Decullier E, Moreau T, Chapuis F, Onen F. Observation-based nocturnal sleep inventory: screening tool for sleep apnea in elderly people. J Am Geriatr Soc 2008;56:1920–5.
19. Vennelle M, Engleman HM, Douglas NJ. Sleepiness and sleep-related accidents in commercial bus drivers. Sleep Breath 2010;14:39–42.
20. Whitney CW, Enright PL, Newman AB, Bonekat W, Foley D, Quan SF. Correlates of daytime sleepiness in 4578 elderly persons: the Cardiovascular Health Study. Sleep 1998;21:27–36.
21. Ensrud KE, Blackwell TL, Redline S, et al. Sleep disturbances and frailty status in older community-dwelling men. J Am Geriatr Soc 2009;57:2085–93.
22. Johansson P, Alehagen U, Svanborg E, Dahlstrom U, Brostrom A. Sleep disordered breathing in an elderly community-living population: relationship to cardiac function, insomnia symptoms and daytime sleepiness. Sleep Med 2009;10:1005–11.
23. Foley DJ, Masaki K, White L, Larkin EK, Monjan A, Redline S. Sleep-disordered breathing and cognitive impairment in elderly Japanese-American men. Sleep 2003;26:596–9.
24. Practice parameters for the indications for polysomnography and related procedures. Polysomnography Task Force, American Sleep Disorders Association Standards of Practice Committee. Sleep 1997;20:406–22.
25. Kushida CA, Littner MR, Morgenthaler T, et al. Practice Parameters for the indications for polysomnography and related procedures: an update for 2005. Sleep 2005;28:499–521.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication July, 2015
Submitted in final revised form October, 2015
Accepted for publication November, 2015
Address correspondence to: S.-Hakki Onen MD, PhD, Hôpital Eduard Herriot, Geriatric Sleep Medicine Center, 5 Place d'Arsonval, 69003 Lyon, France; Tel: +33 4 72 11 11 56 ; Fax: +33 4 27 85 80 90 ; Email address: saban-hakki.onen@chu-lyon.fr

DISCLOSURE STATEMENT

This was not an industry supported study. The authors have indicated no financial conflicts of interest. Drs F. Onen and S.H Onen own the copyright to the ODSI (Copyright © Onen & Onen). Drs F. Onen and S.H Onen permit use of the ODSI by individual people (including clinicians and researchers) free of charge. All other users of the ODSI, such as Corporations, Governmental Agencies, and funded groups (e.g., group practices, hospitals, not-for-profit organizations, University institutions) must have a license to use the ODSI. All enquiries about using the ODSI, whether free of charge or under license, should be made to Dr S.H. Onen saban-hakki.onen@chu-lyon.fr