The Barcelona Sleepiness Index: A New Instrument to Assess Excessive Daytime Sleepiness in Sleep Disordered Breathing

Marc Guaita, DMD[1,6]; Manel Salamero, MD, PhD[1,2,6,10]; Isabel Vilaseca, MD, PhD[1,3,7,10]; Alex Iranzo, MD, PhD[1,4,6,8];
Josep M. Montserrat, MD, PhD[1,5,7,10]; Carles Gaig, MD, PhD[1,4,8]; Cristina Embid, MD, PhD[1,5,7]; Montserrat Romero, BSc[1,6];
Mònica Serradell, BSc[1,9]; Carme León, RN[1,5]; Joan de Pablo, MD, PhD[1,2,6,10]; Joan Santamaria, MD, PhD[1,4,6,8,10]

[1]Multidisciplinary Sleep Disorders Unit, Hospital Clínic of Barcelona, Barcelona, Spain; [2]Department of Psychiatry, Hospital Clínic of Barcelona, Barcelona, Spain; [3]Department of Otorhinolaryngology, Hospital Clínic of Barcelona, Barcelona, Spain; [4]Department of Neurology, Hospital Clínic of Barcelona, Barcelona, Spain; [5]Department of Pneumology, Hospital Clínic of Barcelona, Barcelona, Spain; [6]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain; [7]Ciber Enfermedades Respiratorias (CIBERES), Madrid, Spain; [8]Ciber Enfermedades Neurológicas (CIBERNED), Madrid, Spain; [9]Fundació Clínic per a la Recerca Biomèdica, Barcelona, Spain; [10]University of Barcelona Medical School, Barcelona, Spain

Study Objectives: To develop the Barcelona Sleepiness Index (BSI), an interviewer-administered instrument for assessing excessive daytime sleepiness (EDS) in sleep-disordered breathing (SDB) that correlates well with objective measures of EDS and which is sensitive to change with treatment.

Methods: (1) Generation of a preliminary item list: Fifty-three consecutive SDB patients complaining of EDS and their bed partners were interviewed using a focus group methodology to generate a list of situations prone to cause sleepiness. Sixty different consecutive SDB patients were then evaluated using cognitive interviews to refine this list. (2) Construct validity: The maintenance of wakefulness test (MWT), the multiple sleep latency test (MSLT) and the sustained attention to response task (SART) test were used in an additional 98 consecutive SDB patients with and without EDS. The item combination that best correlated with the objective tests constituted the BSI. Cutoff values were determined to differentiate between patients with and without EDS. (3) Sensitivity to change: Thirty patients requiring continuous positive airway pressure (CPAP) were evaluated after satisfactory treatment.

Results: A combination of two items, "in the morning, when

relaxing" and "in the afternoon, standing inactive, in a public place," presented the highest correlations with the MWT (r: −0.50, p < 0.001), the MSLT (r: −0.21, p = 0.07), and the SART (r: 0.27, p < 0.02) and constituted the BSI. The BSI significantly correlated with oxyhemoglobin saturation measures (nadir SpO$_2$: r: −0.28, p = 0.01; CT 85: r: 0.23, p = 0.04) and showed a high sensitivity to change with CPAP treatment (t: 3.4, p = 0.002). A score of 2 or more identified patients with objective EDS (sensitivity = 64.9%, specificity = 72.1%).

Conclusion: The Barcelona Sleepiness Index is a simple, brief instrument for measuring subjective EDS in SDB.

Keywords: Barcelona Sleepiness Index, excessive daytime sleepiness, focus group, MSLT, MWT, SART, sleep disordered breathing

Citation: Guaita M, Salamero M, Vilaseca I, Iranzo A, Montserrat JM, Gaig C, Embid C, Romero M, Serradell M, León C, de Pablo J, Santamaria J. The Barcelona Sleepiness Index: a new instrument to assess excessive daytime sleepiness in sleep disordered breathing. *J Clin Sleep Med* 2015;11(11):1289–1298.

Sleepiness is a physiological phenomenon that becomes a problem when manifest under abnormal circumstances that interfere with daytime activities. One of the most common causes of excessive daytime sleepiness (EDS) is sleep disordered breathing (SDB), a condition associated with driving accidents, psychosocial morbidity, cardiovascular risk, and poor quality of life.[1,2] Assessment of EDS is an important part of the evaluation and management of SDB patients and a key decision point in the treatment algorithm.

EDS is usually evaluated objectively, employing direct electrophysiological recordings or indirect behavioral measures, and subjectively, using sleepiness scales. The multiple sleep latency test (MSLT) and the maintenance of wakefulness test (MWT) are the two most commonly used objective tools for characterizing the ability and the resistance to fall asleep, respectively.[3,4] However, they are relatively complex, require trained technicians to interpret the signals, and are expensive


BRIEF SUMMARY

Current Knowledge/Study Rationale: Subjective sleepiness scales have weak correlations with objective measures of sleepiness and disease severity in patients with sleep disordered breathing. A new questionnaire developed using focus group techniques and validated with different objective sleepiness measures could be useful in clinical practice.

Study Impact: The Barcelona Sleepiness Index is a brief questionnaire of just two items, which correlates well with objective sleepiness measures, oxyhemoglobin desaturation and is sensitive to change with therapy. This instrument could be helpful in the evaluation of sleepiness, both during routine clinical interviews as well as a screening method in epidemiological studies.

---

to perform on a daily basis. Behavioral performance tests, such as reaction time tests—psychomotor vigilance test (PVT) or the sustained attention to response task (SART)—have been used as alternative tools to measure decrements in vigilant attention associated with sleepiness.[5,6] However, performance

Figure 1—Development and validation of the Barcelona Sleepiness Index: study design.
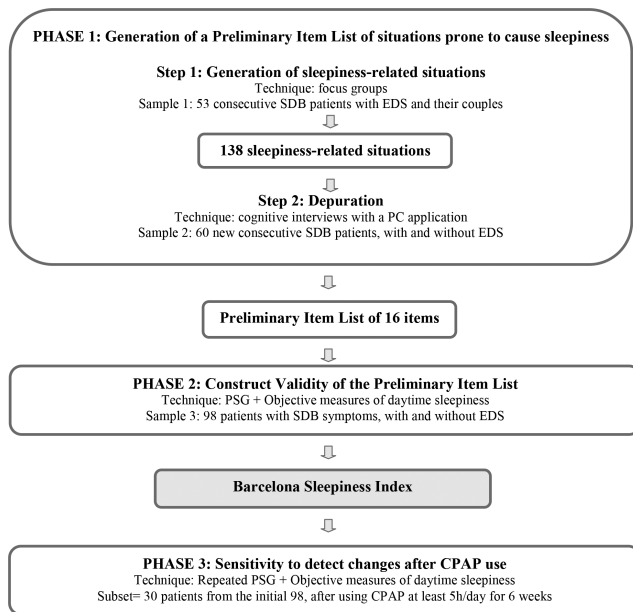
SDB, sleep disordered breathing; EDS, excessive daytime sleepiness; PSG, polysomnography; CPAP, continuous positive airway pressure.

does not always correlate with the most widely accepted sleepiness measures,[7,8] because it is influenced by other factors such as task duration, motivation, and complexity.[9]

Three self-report sleepiness scales have been developed and validated relative to the MSLT in adults, the Epworth Sleepiness Scale (ESS),[10] the Sleep-Wake Activity Inventory (SWAI),[11] and the Sleepiness-Wakefulness Inability and Fatigue Test (SWIFT).[12] The ESS asks subjects to rate, in relation to "recent times," their probability of dozing off or falling asleep in 8 different situations commonly encountered in daily life. The SWAI is a 6-subscale questionnaire of 59 items that includes a 9-item Daytime Sleepiness Subscale. In this subscale, subjects rate how often each sleepiness-related item occurred during the preceding week. In contrast, the SWIFT is a 12-item scale evaluating fatigue and the inability to maintain wakefulness during the preceding month in situations where staying awake is desirable.

The ESS, unlike the SWAI and SWIFT, has been widely used in both clinical and research settings and is often considered the scale of reference. It is easy to fill out, differentiates between different levels of EDS, and is sensitive to treatment-induced changes. However, it has a weak correlation with objective sleepiness tests and with SDB nocturnal parameters,[13,14] and there is frequent disagreement between patient and partner scores of the patient's sleepiness.[15] In clinical practice, these limitations could increase the risk of misclassification and inadequate treatments.

The design of the ESS has been criticized on the grounds that two items are virtually identical ("Sitting quietly after lunch without alcohol" and "Lying down to rest in the afternoon when circumstances permit"), two are overly general ("sitting and reading" and "watching TV"),[16] and one is ambiguous ("in a car, while stopped for a few minutes in traffic").[13] Indeed, how the number and type of items that make up the ESS were

chosen has never been clearly explained.[10,13] In recent decades, questionnaire design has improved thanks to the development of such techniques as the focus group. The latter allows the real concerns of a specific group of people to be identified, takes into account the actual language and expressions used by these people, and reflects the consensus among the members of that group.[17,18] However, no instruments or scales using this approach have yet been developed to measure EDS.

In addition, the validation of the existing ESS, SWAI, and SWIFT have relied solely on the MSLT and have overlooked all other objective measures available (e.g., the MWT or behavioral performance tests).[10–12] Yet, given that each of these tests seems to capture different aspects of sleepiness,[19] the combined use of their results might provide a better reflection of an individual's global level of EDS.

For these reasons, we decided to develop and validate a new instrument to measure EDS in SDB using the best methodology available. Our goal was to design an instrument that reflected the sleepiness complaints of the patients with the highest possible correlation with objective sleepiness tests (MSLT, MWT, and SART) and one that was sensitive to detect changes after adequate continuous positive airway pressure (CPAP) treatment.

METHODS

Study Design

The study was conducted in the Multidisciplinary Sleep Disorders Unit in the Hospital Clinic, Barcelona (Spain), a tertiary university hospital. The study was approved by the local ethics committee. All subjects received written information and signed informed consent prior to participation. An overview of the development and validation of the Barcelona Sleepiness Index (BSI) is shown in **Figure 1**. The design of the instrument included the following steps: (1) generation of a preliminary item list (PIL) of situations that may cause sleepiness, (2) construct validity using objective measures of EDS, and (3) assessment of sensitivity to change after adequate CPAP therapy.

Phase 1: Generation of a Preliminary Item List of Situations Prone to Cause Sleepiness

Fifty-three consecutive SDB patients complaining of EDS and their bed partners were interviewed in groups of 3 to 5 couples using focus group techniques. A psychologist conducted the session that focused on daily situations that may be prone to give rise to sleepiness. Each session lasted 2 h and was recorded and transcribed. The textual analysis of the transcripts yielded 243 EDS situations: 156 in which the patient could actually "fall asleep" and 87 in which the patient perceived "sleepiness without falling asleep." Experts from our Unit reviewed these 243 situations and distilled a list of 25 prototypical situations with 4 modulators: time of day, body position, motivation, and duration of the activity (**Table S1**, supplemental material). The logical combination of the situations with the modulators yielded a list of 138 sleepiness-related situations.

A new group of 60 consecutive SDB patients with and without complaints of EDS responded to the list of 138 situations. For each situation, they were asked to answer (1) how severe

the sleepiness they felt was (1 = "not feeling sleepy," 2 = "feeling sleepy without falling asleep," 3 = "feeling sleepy and falling asleep," 4 = "falling asleep unexpectedly"), (2) how often sleepiness occurred (1 = "sometimes," 2 = "often," 3 = "always or most of the time"), and (3) how much time would elapse before they fell asleep while engaged in the situation (3 = "less than 5 minutes," 2 = "between 5 and 15 minutes," or 1 = "more than 15 minutes"). A total score for each situation was computed by multiplying the severity of sleepiness, the frequency of occurrence and the time needed to fall asleep, with values ranging from 1 to 36. The mean scores of the 138 situations were logarithmically transformed to correct for the skewness of the distribution and plotted in order of magnitude. On this basis, we selected the situations that were identified with greatest frequency and which provided a homogeneous representation of the continuum of the severity of sleepiness. In case of ties or close proximity, the tie with the greatest frequency of endorsement was selected. As a result, a PIL of sleepiness-related situations, expressed in Spanish, was used for the validation study.

The PIL was formatted prior to the evaluation of subjective EDS using an interviewer-administered approach. The items were grouped according to time of day and degree of physical/mental activity and both the intensity and the circadian pattern of sleepiness were recorded. Patients were asked to report if during "recent weeks" they had felt sleepy and/or fallen asleep when in each situation, their responses being categorized according to the severity of sleepiness, the frequency of occurrence and the time elapsed before falling asleep. Three possible scoring systems were considered for the analysis: (a) severity "alone," (b) severity weighted by the frequency of occurrence, and (c) severity weighted by the frequency of occurrence and the sleep latency.

Phase 2: Construct Validity Using Objective Measures of EDS

Patients

A new cohort of 98 consecutive patients complaining of snoring or apneas with and without EDS was evaluated. Exclusion criteria were age under 18 years, major medical or psychiatric disorders, use of medications affecting wakefulness or sleep, shifts, and irregular sleep-wake schedules the week before the sleep study, as shown by actigraphy.

Procedure

Patients underwent a 24-h sleep study to diagnose SDB (nocturnal study) and to assess EDS with objective tests (daytime studies) (**Table 1**). The patients' sleep-wake schedules in the week prior to laboratory sleep study were monitored by actigraphy and using sleep diaries in line with the methodology described by Ortiz et al.[20]

Questionnaires

On admission, subjective daytime sleepiness and mood disorder symptoms were evaluated with the Spanish version of the ESS and the Hospital Anxiety and Depression Scale (HADS), respectively.[21] These 2 scales were self-administered. The PIL of sleepiness-related situations created by our group was then administered by one of the authors (MG) to ensure it was properly understood.

Table 1—Construct validity: schedule.

18:00	Admission to the sleep lab
18:30	ESS, HAD, PIL of 16 items
19:00	Electrode placement
20:00	Dinner
23:00	Start PSG
07:30	End PSG
07:30–08:15	Breakfast
08:15–08:20	1st SART
08:30–09:10	1st MWT
09:30–09:50	1st MSLT
10:15–10:20	2nd SART
10:30–11:10	2nd MWT
11:30–11:50	2nd MSLT
12:15–12:20	3rd SART
12:30–13:10	3rd MWT
13:30–13:50	3rd MSLT
13:50–14:15	Lunch
14:15–14:20	4th SART
14:30–15:10	4th MWT
15:30–15:50	4th MSLT
16:15–16:20	5th SART
16:30–17:10	5th MWT
17:30–17:50	5th MSLT

ESS, Epworth Sleepiness Scale; HAD, Hospital Anxiety and Depression Scale; PIL, preliminary item list; PSG, polysomnography; MWT, maintenance of wakefulness test; MSLT, multiple sleep latency test; SART, sustained attention to response task.

Nocturnal Polysomnography

Polysomnography (PSG) was performed from 23:00 until 07:30 with a digital polygraph (Deltamed, software version 2007, Paris, France) to confirm SDB and to rule out any other significant comorbid sleep disorders (e.g., REM sleep behavior disorder, epileptiform EEG activity). PSG included electroencephalography (O2-A1, O1-A2, C4-A1, C3-A2, F4-A1, F3-A2), electrooculography, electrocardiography, submentalis and right and left anterior tibialis surface electromyography, and synchronized audiovisual recording. Nasal cannulae, nasal and oral thermistors, abdominal and thoracic strain gauges, and finger pulse oximeters were used to measure respiratory variables. Apnea was defined as a complete cessation of airflow for ≥ 10 sec using thermistor signal. Hypopnea was defined as ≥ 30% reduction in nasal pressure signal excursions from baseline and associated with ≥ 3% desaturation from pre-event baseline or arousal. The apnea-hypopnea index (AHI) was the number of apneas plus hypopneas per hour of sleep. An AHI > 5 was considered indicative of SDB. Sleep stages were scored manually according to the American Academy of Sleep Medicine (AASM) criteria using 30-s epochs.[22]

Objective Measures of EDS

The morning following the nocturnal PSG, the patients initiated a protocol to measure EDS objectively throughout the day (**Table 1**). This protocol comprised 5 blocks of MWT followed by MSLT (research version),[23] every 2 h starting at 08:30. The order of the tests was the same for all subjects and the setting

conditions, light intensity and temperature adhered to standard recommendations.[4] Each nap block was preceded by a measurement of vigilance with the SART. The SART lasted 4 min, was performed in a dark, quiet room with the patient seated in front of a computer screen, and followed the recommendations of Robertson et al.[6] MWT and MSLT naps ended after 40 and 20 min, respectively, if no sleep occurred or immediately after unequivocal sleep, defined as 3 consecutive epochs of stage N1 sleep, or one epoch of any other stage of sleep. Mean sleep latency to the occurrence of the first of 3 consecutive epochs of stage N1 sleep or any other single sleep stage epoch was used as the operational measure of objective sleepiness for the MSLT and MWT,[3] while the number of "commission errors" was used for the SART (key presses when no key should be pressed, i.e., after a 3).[6]

Phase 3: Assessment of Sensitivity to Change after CPAP Therapy

Thirty patients from the original cohort of 98 were treated with CPAP and were evaluated again following the same protocol after optimal therapy. CPAP titration was performed following the recommendations of the Spanish Sleep Network.[24] CPAP compliance was measured objectively using a built-in CPAP meter. A minimum CPAP use of 5 h per night during 6 consecutive weeks was required for the analysis.[25] CPAP was used during the nocturnal PSG but not during the MSLT and MWT.

Statistical Analyses

We explored the correlation between the different measures of objective EDS and the scores of each PIL item obtained using the three scoring systems (severity "alone," severity weighted by the frequency of occurrence, and severity weighted by the frequency of occurrence and sleep latency). The objective test and the scoring system with the highest overall correlations were selected, the former as the main indicator of objective sleepiness and the latter as the best subjective scoring system. Exhaustive regression analyses using adjusted r^2 as the statistical criterion were performed to find the item or combination of items most closely associated with the selected objective test. The items selected formed the Barcelona Sleepiness Index (BSI).

The total score on the BSI was computed by adding the selected items. Pearson correlation coefficients were calculated between the BSI score and the objective test measurements and the sleep-related respiratory parameters and the same was done for the ESS score. The correlations of the BSI and the ESS with the other variables were compared using Student t-test. Receiver operating characteristic (ROC) curves and optimal cutoff scores were determined by maximizing sensitivity and specificity simultaneously. Finally, the BSI scores before and after CPAP treatment were compared with paired Student t-tests. The significance level was set at $p < 0.05$. All statistical analyses were performed with R version 3.1.

RESULTS

Phase 1: Generation of a Preliminary Item List of Situations Prone to Cause Sleepiness

The logarithm of the mean scores for each of 138 situations prone to cause sleepiness is plotted in order of magnitude in

Figure 2. We selected 14 situations that show a high frequency of response and which represent homogenously the sleepiness severity continuum. Two additional situations related to sleepiness while driving (items 7 and 10) were added because of the clinical relevance of car accidents. Thus, a PIL of 16 items was used for the validation study of the BSI (**Table S2**, supplemental material).

Phase 2: Construct Validity

Clinical and PSG Characteristics

Of the 98 consecutive patients originally studied, 9 were excluded due to irregular sleep-wake rhythms (3 patients), acute sleep deprivation prior to the sleep study (1 patient), technical problems during the procedure (3 patients), severe depressive symptoms (1 patient) and REM sleep without atonia (1 patient). The validity group finally comprised 89 adults with suspected SDB, and their clinical and PSG characteristics are shown in **Table 2**. In brief, there was a wide spectrum of disease severity ranging from simple snoring (AHI < 5: 12.4%) to mild (AHI 5–15: 33.3%), moderate (AHI 15–30: 22.5%), and severe SDB (AHI > 30: 34.8%). Subjective and objective measures of daytime sleepiness demonstrated a wide range of daytime sleepiness accompanying SDB (**Table 3**).

#### Frequency of Responses to the Preliminary Item List

All but 3 items (9, 11, and 12) were identified by > 70% of the patients and so represented frequent situations in which to evaluate EDS. Most patients did not feel sleepy while standing (items 4, 5, 6, 14, and 15), sitting on the toilet (item 2) or having dinner (item 16), while 99% of respondents reported feelings of sleepiness and/or falling asleep when watching TV with little interest after lunch (item 8) (**Figure 3**). In general, increased severity of sleepiness was reported in situations requiring a low level of attention or motivation (items 1, 3, 9, 11, 12, and 13). Driving was identified as a cause of sleepiness in 16% of patients in the morning (item 7) and in 31% of patients after lunch (item 10). Finally, in terms of the time of day, EDS was most frequently reported after lunch.

Relation with Objective Measures of EDS

The best correlation between the 16 items and the objective measures of sleepiness was found for the scoring system that considered "severity alone," as shown in **Table 4**. Weighting severity by frequency of occurrence and sleep latency worsened these correlations and so these scoring systems were not used any further in the analysis. Correlation coefficients were higher with the MWT than they were with the MSLT or the SART, while the mean "committed errors" on the SART correlated better with the 16-item scores than they did with the MSLT, which presented lower, more heterogeneous and sometimes contradictory correlations. (**Table 4**) Therefore, we concluded that the MWT was the main objective criteria for measuring EDS.

The Barcelona Sleepiness Index (BSI)

After regression analysis, we found that a combination of 2 items, "in the morning, when relaxing" (item 1) and "in the afternoon, when standing, inactive in a public place" (item 14), achieved the highest predictive value of the MWT sleep latencies (multiple r = −0.51, p < 0.001) and constituted the

Figure 2—Preliminary item list: item selection.

(A)
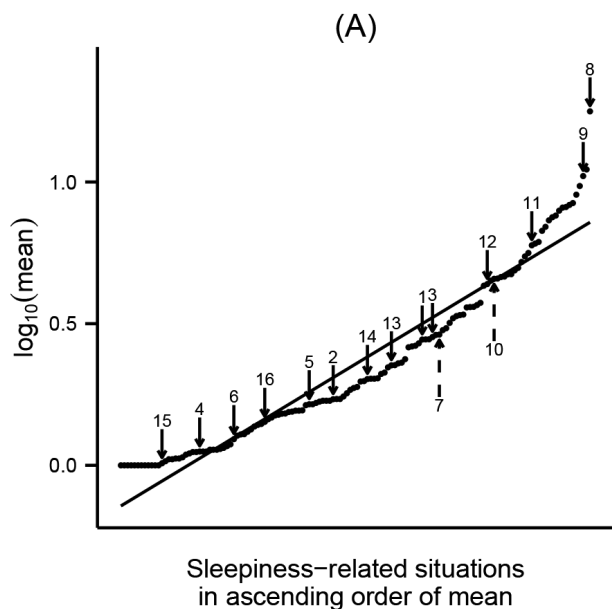
(B)

In the morning:
1. When relaxing
2. When sitting on the toilet
3. When at a boring get-together with family or friends
4. When standing still
5. When standing on public transport
6. When standing and speaking to someone with little interest
7. When driving and you are stopped at the lights

After lunch:
8. When watching TV with little interest
9. When sitting in a waiting room
10. When driving and you are stopped at the lights

In the afternoon:
11. When sitting down watching a boring show (theatre, cinema, etc.)
12. When you are in the passenger seat of a car on an interesting trip
13. When sitting down and speaking to someone with little interest
14. When standing, inactive in a public place (waiting to be served, or for the bus, subway, or a friend, etc.)
15. When standing, speaking to someone on the phone with interest

At night:
16. When having supper

(A) The chart shows the mean score for each of 138 sleepiness-related situations plotted in order of magnitude. A logarithmic transformation was performed to correct for the skewness and to represent a linear sleepiness severity continuum (regression line). Fourteen items (solid arrows) distributed homogeneously along the graph were identified by most of the subjects interviewed and were selected. Two additional situations related to sleepiness during driving (dashed arrows) were added because of the clinical relevance of car accidents. The preliminary item list is shown in (B).

Barcelona Sleepiness Index (BSI). The inclusion of additional items did not improve this predictive value (Δr p > 0.05). The original Spanish version of the BSI and its English translation are shown in **Tables 5** and **6**.

The BSI total score is the sum of the 2 item scores and ranges between 0 and 6. In our sample, total scores ranged from 0 to 3 and were affected mainly by the scoring on the passive item—in the morning, when relaxing. For this item, 40% of the patients reported "feeling sleepy and falling asleep" (score = 2)

Table 2—Clinical and PSG characteristics of the study sample.

	Results
Clinical data	
Male (%)	74
Age (years old)	52.9 ± 11.1
Body Mass Index (kg/m²)	29.6 ± 5.3
HADS-A	6.2 ± 3.6
HADS-D	3.7 ± 3.1
Actigraphy 7 days	
Mean nocturnal TST (min)	427.4 ± 57.5
Polysomnographic data	
TIB (min)	470.0 ± 25.1
TST (min)	388.6 ± 53.9
SE (%)	82.5 ± 10.5
WASO (min)	64.7 ± 42.1
Stage 2 sleep latency (min)	17.8 ± 16.5
REM sleep latency (min)	110.2 ± 70.2
Stage 1 (%)	17.4 ± 10.4
Stage 2 (%)	56.1 ± 8.8
Stage 3 (%)	10.5 ± 7.6
REM sleep (%)	16.0 ± 5.8
REM stages (n)	3.7 ± 1.4
Sleep stage changes (n)	175.2 ± 61.7
Periodic Limb Movement Index (events/h)	9.4 ± 17.0
Arousal Index (events/h)	32.5 ± 19.5
AHI (events/h)	28.8 ± 25.5
Mean SpO_2 (%)	93.4 ± 2.7
CT90 (%)	10.2 ± 18.2
Oxygen Desaturation Index 3% (events/h)	22.5 ± 24.6
Nadir SpO_2	81.0 ± 9.5

All values are mean ± standard deviation. PSG, polysomnography; HADS-A, Hospital Anxiety and Depression Scale - Anxiety; HADS-D, Hospital Anxiety and Depression Scale - Depression; TIB, time in bed; TST, total sleep time; SE, sleep efficiency; WASO, wake after sleep onset; AHI, apnea-hypopnea index; mean SpO_2, mean O_2 saturation; CT90, cumulative time spend with an O_2 saturation below 90%; Nadir SpO_2, minimal O_2 saturation value; n, number.

whereas only 10% of the patients reported "feeling sleepy but not falling asleep" (score = 1) in the afternoon, when standing, inactive in a public place (see the frequencies of response to item 1 and item 14 in **Figure 2**).

Comparison between the BSI and the ESS: Correlation Analysis

The BSI presented stronger correlations with the 3 objective tests than it did with the ESS: MWT ($r_{BSI} = -0.50$ *vs* $r_{ESS} = -0.41$), MSLT ($r_{BSI} = -0.21$ *vs* $r_{ESS} = 0.02$), and SART ($r_{BSI} = 0.27$ *vs* $r_{ESS} = 0.20$). The differences between scales, however, did not achieve statistical significance, although the BSI tended to better correlate with the MSLT than the ESS (see **Table 7**). Finally, the BSI was significantly correlated with the ESS (rho = 0.52, p < 0.001).

In contrast to the ESS, the BSI score was significantly correlated with the oxyhemoglobin saturation measures (nadir SpO_2: $r_{BSI} = -0.28$ *vs* $r_{ESS} = 0.07$, p < 0.01) and with the CT 85

(r_{BSI} = 0.23 *vs* r_{ESS} = −0.01, p = 0.05) (**Table 7**). Neither the BSI nor the ESS was correlated with the arousal index or with the percentage of Stage N3 sleep.

The BSI: Cutoff Values for Objective EDS

The area under the curve (AUC) between patients with MWT < 20 min and MWT ≥ 20 min was 0.72 (95CI: 0.61–0.82) for the BSI and 0.68 (95CI: 0.57–0.79) for the ESS, with no significant difference between the 2 instruments (p = 0.89). A cutoff score was calculated by simultaneously optimizing sensitivity and specificity. A cutoff of 2 for the BSI showed a sensitivity of 64.9% and a specificity of 72.1% for discriminating between the sleepy (< 20 min) and non-sleepy (≥ 20 min) groups (**Table 8**). The same procedure gave a cutoff of 13 for the ESS, with sensitivity of 56.4% and a specificity of 61.2%. The commonly used cutoff ≥ 11 for the ESS achieved in our sample gave a sensitivity

of 69.2% and a specificity of 46.9%. MWT mean sleep latencies and BSI scores are represented in **Figure 4**.

Figure 3—Preliminary item list: frequency of response to each item.
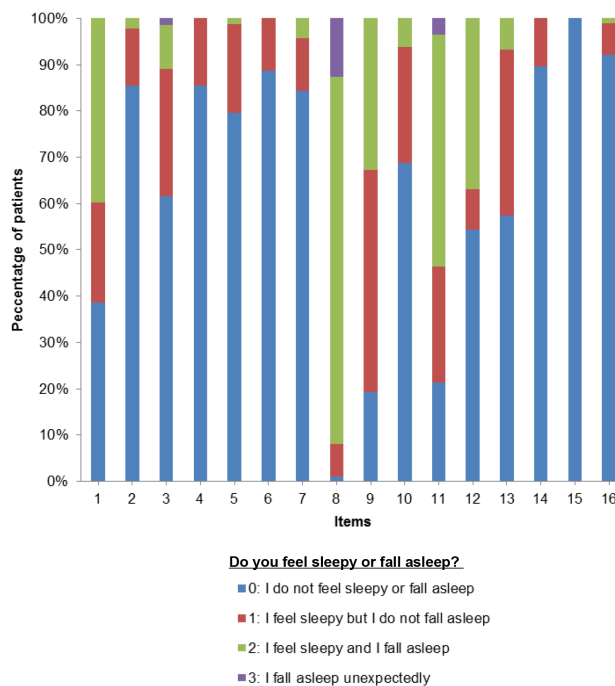
Percentage of patients (vertical axis) responding to each item of the preliminary item list (horizontal axis). Scores are represented with different colors and ranged from minimal (score = 0) to maximal EDS (score = 3). Items are ordered according to the time of day (see Table S2 for detail). EDS, excessive daytime sleepiness.

Table 3—Daytime sleepiness measures of the study sample.

	Mean ± SD	Range	% with EDS
ESS score (n)	11.7 ± 4.3	4–22	60.7[a]
MWT (min)	24.3 ± 11.8	4.7–40	44.3[b]
MSLT (min)	7.9 ± 4.2	0.9–18.2	56.6[c]
SART (p.e)	7.0 ± 4.7	0.6–21.6	n.a

Cut-off values of excessive daytime sleepiness: [a]ESS > 10, [b]MWT < 20 min, [c]MSLT < 8 min. Normative values for SART are not available (n.a.). ESS, Epworth Sleepiness Scale; MWT, maintenance of wakefulness test; MSLT, multiple sleep latency test; SART, Sustained Attention to Response Task; EDS, excessive daytime sleepiness; min, minutes; p.e, pressed errors; n, number.

Table 4—Correlations between the severity score of the PIL items and objective measures of EDS.

	MWT	MSLT	SART
In the morning:			
1. When relaxing	−0.43***	−0.15	0.20
2. When sitting on the toilet	−0.09	0.21	−0.09
3. When at a boring get-together with family or friends	−0.23*	0.12	0.18
4. When standing still	−0.25*	−0.12	0.04
5. When standing on public transport	−0.27*	−0.12	0.10
6. When standing and speaking to someone with little interest	−0.16	0.09	−0.01
7. When driving and you are stopped at the lights	−0.29*	−0.19	0.09
After lunch:			
8. When watching TV with little interest	−0.11	0.10	0.29*
9. When sitting in a waiting room	−0.33*	0.05	0.21
10. When driving and you are stopped at the lights	−0.26*	0.01	0.12
In the afternoon:			
11. When sitting down watching a boring show (theatre, cinema, etc.)	−0.12	0.14	0.46**
12. When you are in the passenger seat of a car on an interesting trip	−0.23	0.06	−0.02
13. When sitting down and speaking to someone with little interest	−0.21	0.11	0.21
14. When standing, inactive in a public place (waiting to be served, or for the bus, subway, or a friend, etc.)	−0.25*	−0.23*	0.21
15. When standing, speaking to someone on the phone with interest	–	–	–
At night:			
16. When having supper	−0.20	−0.02	0.01

*p < 0.05; **p < 0.01; ***p < 0.001. No results are shown for item 15 because no patient experienced sleepiness in that situation. PIL, preliminary item list; EDS, excessive daytime sleepiness; MSLT, multiple sleep latency test; MWT, maintenance of wakefulness test; SART, sustained attention to response task.

Table 5—Barcelona Sleepiness Index: Spanish version.

Instrucción al CLÍNICO:

1) Estas preguntas se refieren a la somnolencia del paciente **durante las últimas semanas**.

2) En caso que el paciente refiera inicialmente no experimentar estas situaciones, ayudarlo poniendo ejemplos.

¿Presenta somnolencia o se duerme…?	NO	SÍ, tengo sueño PERO no me duermo	SÍ, tengo sueño Y me duermo	SÍ, me duermo sin darme cuenta
1. Por la mañana, cuando se relaja	0	1	2	3
2. Por la tarde, de pie inactivo, en un lugar público (haciendo cola, esperando a un amigo, el autobús, el metro…)	0	1	2	3

Puntuación total (ítem 1+ ítem 2) = _____

Una puntuación total ≥ 2 indica somnolencia diurna excesiva.

Table 6—The Barcelona Sleepiness Index: English version.

Instruction to the CLINICIAN

1) These questions refer to the sleepiness experienced by the patient **during the recent weeks**.

2) If the patient initially denies experiencing these situations, please give more examples.

Do you feel sleepiness or fall asleep…?	NO	YES, I feel sleepy BUT I do not fall asleep	YES, I feel sleepy AND I fall asleep	YES, I fall asleep unexpectedly
1. In the morning, when relaxing	0	1	2	3
2. In the afternoon, when standing, inactive in a public place (waiting to be served, or for the bus, subway, or a friend, etc.)	0	1	2	3

BSI total score (item 1+ item 2) = _____

If score ≥ 2 points, it is indicative of excessive daytime sleepiness.

Table 7—Comparison between the BSI and ESS: correlation analysis.

	BSI	ESS	Correlation Differences (p)
MWT (min)	−0.50***	−0.41***	0.632
MSLT (min)	−0.21	0.02	0.077
SART (p.e)	0.27*	0.20	0.775
Nadir SpO$_2$ (%)	−0.28*	0.07	< 0.01
CT 85 (%)	0.23*	−0.01	0.055

*p < 0.05, **p < 0.01, ***p < 0.001. BSI, Barcelona Sleepiness Index; ESS, Epworth Sleepiness Scale; MWT, maintenance of wakefulness test; MSLT, multiple sleep latency test; SART, sustained attention to response task; Nadir SpO$_2$, minimal O$_2$ saturation value; CT85, cumulative time spend with a O$_2$ saturation below 85%; min, minutes; p.e, pressed errors.

Table 8—The BSI: cutoff scores for detecting objective EDS.

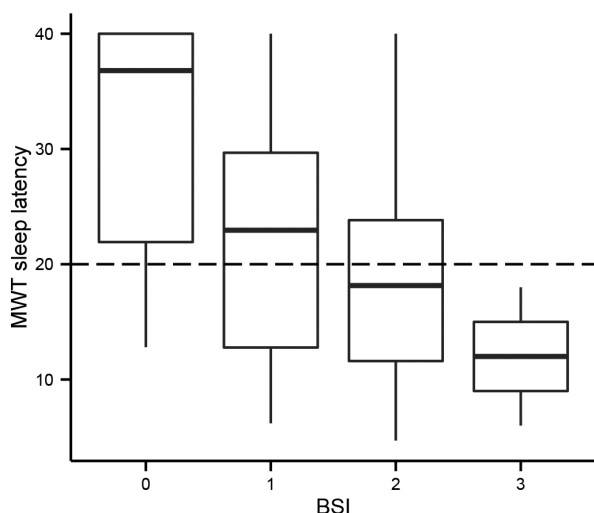|	Cut-Off Scores	Specificity	Sensitivity
BSI	2	72.1%	64.9%
ESS	13	61.2%	56.4%
ESS	11	46.9%	69.2%

BSI, Barcelona Sleepiness Index; ESS, Epworth Sleepiness Scale; EDS, excessive daytime sleepiness.

The BSI score showed significant changes with CPAP treatment (p = 0.002), in parallel with the MWT (p < 0.001), MSLT mean sleep latencies (p = 0.01), and the ESS (p < 0.001). In contrast, SART key press errors did not vary with the treatment. Changes in sleepiness measures are shown in **Table 9**.

DISCUSSION

In the present study, we have shown that a brief instrument of just two items, the Barcelona Sleepiness Index (BSI), presents high correlations with objective measures of sleepiness and is sufficiently sensitive to detect changes after CPAP therapy in SDB patients. This simple interviewer-administered

Phase 3: Sensitivity to Change with CPAP Treatment

One patient successfully treated with CPAP refused to complete the protocol and was withdrawn. Twenty-nine moderate-to-severe SDB patients (24 males, baseline AHI 52.5 ± 23.6) with a minimal compliance of 5 h/day were re-evaluated 192.9 ± 107.1 days after the baseline study. PSG revealed a complete resolution of SDB in all patients. Differences in clinical and PSG characteristics are shown in **Table S3** (supplemental material).

Figure 4—BSI scores and MWT sleep latencies.

The BSI total score is plotted on the horizontal axis and the MWT mean sleep latency on the vertical axis. At baseline, BSI scores ranged between 0 and 3 (with a maximum of 6 points, not shown on the axis). The horizontal dashed line is drawn at the MWT = 20 min, which represents the boundary between patients with and without objective daytime sleepiness. Note that BSI scores increased as MWT sleep latencies became shorter. MWT, maintenance of wakefulness test; BSI, Barcelona Sleepiness Index.

Table 9—Changes in daytime sleepiness measures after CPAP treatment.

	Baseline	CPAP	p value
BSI score (n)	1.5 ± 0.9	0.7 ± 0.9	0.002
ESS score (n)	12.4 ± 3.9	7.4 ± 4.1	< 0.001
MWT (min)	21.8 ± 11.7	31.4 ± 11.1	< 0.001
MSLT (min)	7.1 ± 3.4	9.3 ± 4.4	0.012
SART (p.e)	6.9 ± 4.5	6.5 ± 4.8	0.496

BSI, Barcelona Sleepiness Index; ESS, Epworth Sleepiness Scale; MWT, maintenance of wakefulness test; MSLT, multiple sleep latency test; SART, Sustained Attention to Response Task, CPAP, continuous positive airway pressure; min, minutes; p.e, pressed errors; n, number.

instrument, which can be quickly answered, should be of benefit to clinicians in their evaluation of EDS during regular patient visits. Such an approach would allow the interviewer to clarify the items by providing examples where necessary, adapting their language of exposition according to the patient's level of education.[26] Moreover, the interviewer can emphasize the importance of patients basing their responses on events in recent weeks, focusing on actual experiences rather than on hypothetical situations.[27]

The BSI has ecological validity since it captures the way in which patients with SDB report episodes of daytime sleepiness in their own words. In accordance with the latest Food and Drug Administration guidelines for scale development,[27] we used focus groups to create the BSI. This method today represents the gold standard for developing a questionnaire and has not previously been used for generating a scale that evaluates EDS.[10–12,28,29] In fact, the items making up existing sleepiness scales were selected on the basis of physicians' clinical experiences or by reviewing previous subjective scales, but never directly not on patients' reports.

The analysis of focus group discussions involving our patients provided us with information as to how they experienced episodes of daytime sleepiness and which modulators influenced their occurrence. In describing sleepiness, patients drew a clear distinction between "feeling sleepy" and "falling asleep." Thus, we considered this an essential aspect in the evaluation of EDS. In addition, as previously described,[16] contextual factors such as the time of day, body position, degree of motivation, and the duration of the activity were also rated in terms of their influence on sleepiness. We considered these modulators important for enriching the items, facilitating comprehension and avoiding ambiguity. For example, a high percentage of patients reported experiencing EDS when relaxing "after lunch," coinciding with the physiological increase in sleep tendency at this time of day (see item 8 in **Figure 2**). However, this situation does not have the same relevance as falling asleep when relaxing "in the morning," given that episodes of sleepiness should not be expected after a normal night's sleep. A further reason for including the modulators is that they should help in differentiating between low and intermediate levels of EDS.[16]

We used several objective tests to measure the various aspects of sleepiness and to better assess the adequacy of the items. Sangal et al. suggested that the evaluation of a patient with a complaint of sleepiness or lack of alertness may require a battery of tests, not only the MSLT or the MWT, given that each measures only one component of sleepiness.[23] To the best of our knowledge, this is the first study to evaluate simultaneously three measures of sleepiness in SDB, namely the capacity to remain alert and vigilant (SART), the ability to stay awake (MWT), and the ability to fall asleep (MSLT). These tests were chosen to cover the alertness continuum, from high (SART) to low (MSLT) and were repeated five times throughout the day to capture circadian oscillations of sleepiness.[23,30]

Overall we found that the MWT, used individually, presented the highest correlations with the situations prone to cause sleepiness and, as such, the test was considered the main indicator of objective EDS for creating the BSI (see **Table 4**). It has been previously suggested that the MWT, which requires patients to oppose sleepiness in a soporific environment, is better suited to the objective measurement of daytime sleepiness in SDB than is the MSLT.[23,30,31] In fact, the inability to maintain wakefulness in job-related duties, during social interaction, and while driving or working are the reasons why patients seek medical attention, rather than because of an excessive tendency to fall asleep intentionally in appropriate conditions. Thus, the setting of the MWT seems more realistic than that of the MSLT for evaluating daytime sleepiness. Our results support this view because the highest number of significant correlations of the preliminary 16-item list was obtained with the MWT, followed by the SART and then the MSLT.

We found that a combination of two items—a passive situation, "in the morning, when relaxing," and a more active situation, "in the afternoon, when standing inactive, in a public place"—provided the highest predictive values of MWT and

thus constituted the BSI. Although this questionnaire showed similar correlations to those of the ESS with the MWT and the SART, it tended to better correlate with the MSLT ($r_{BSI} = -0.21$ vs $r_{ESS} = 0.02$, p = 0.077) and demonstrated higher correlations with oxyhemoglobin saturation measures. These findings together with the poor correlation between the ESS and the MSLT or with the sleep-related respiratory parameters found in this and previous studies[30,32–34] suggest that the BSI measures daytime sleepiness equally or even better than the ESS in SDB.

The BSI scores also improved after satisfactory CPAP treatment, accompanying the decrease in the ESS scores and the longer sleep latencies recorded on the MWT and MSLT. In line with our findings, other studies have reported an improvement in subjective EDS when using the ESS,[35,36] the daytime sleepiness subscale of the SWAI and the SWIFT.[11,12] In addition, we demonstrated that treatment improved "the ability to stay awake" (MWT) more than the "sleep tendency" (MSLT), in line with the study performed by Sangal et al. in patients with different sleep disorders and therapies.[37] In contrast to these authors, we found a significant lengthening of MSLT sleep latencies with CPAP,[38,39] but not as great as the increments found on the MWT. However, we failed to observe any changes in vigilance and attention with treatment, as measured by the SART. This test has been shown to be abnormal in patients with different causes of EDS including SDB,[8,40] but it has never been used to measure the changes associated with CPAP treatment. Nevertheless, a previous study performed with the PVT failed to find any changes in vigilance after six months of CPAP, thus confirming our results.[41]

The BSI total score ranged from 0 to 3 (with a maximum of 6 points) and did not show a ceiling effect, suggesting that this instrument still has room to detect higher levels of sleepiness than those observed in our patients, which as a group presented moderate levels of sleepiness (see **Table 3**). For example, we might expect BSI scores higher than 3 in SDB patients with extreme levels of EDS or patients with narcolepsy since they might fall asleep unexpectedly in passive situations (3 points on item 1) or while standing up (at least 2 points on item 2). In terms of sensitivity and specificity, we have suggested a tentative cutoff of 2 to identify those patients with and without objective EDS, which yielded higher results than those for the ESS (see **Table 8**).

The design of the preceding study has certain limitations. The construct validity of the BSI has been tested employing a previously unused experimental MSLT/MWT/SART protocol. The duration and the fixed order of the tests may have influenced sleep latencies on the MWT and the MSLT. However, a similar study performed by Sangal et al. demonstrated that the effect of nap-order on sleep latencies was not relevant.[23] A further limitation may lie in the number of patients evaluated after CPAP in order to test the responsiveness of the BSI to this treatment. Finally, we have not assessed the repeatability of BSI scores over a period of time when no real change has occurred. As such, test-retest reliability is needed before the BSI can be definitively recommended.

In conclusion, we have developed an instrument to evaluate EDS in a homogeneous SDB sample with no confounders that might interfere with the main symptom of EDS. It is simple, unambiguous and quick to answer and can be used easily in a traditional face-to-face interview. The psychometric properties of the BSI are certainly promising and should now be confirmed in samples of SDB complaining about fatigue and mood disorder symptoms for the generalization of the results. Since validating a scale is a long process of collecting evidence of its functioning in samples with different sociodemographic, cultural, and clinical characteristics, future research should seek to evaluate the BSI in other cultures and in non-Spanish-speaking populations, include SDB patients with other therapeutic approaches and patients with other causes of EDS (narcolepsy, sleep deprivation syndrome, etc.) to confirm its clinical utility in daily routine practice.

ABBREVIATIONS

AASM, American Academy of Sleep Medicine
AHI, apnea-hypopnea index
BSI, Barcelona Sleepiness Index
CPAP, continuous positive airway pressure
EDS, excessive daytime sleepiness
ESS, Epworth Sleepiness Scale
HAD, hospital anxiety and depression scale
MSLT, multiple sleep latency test
MWT, maintenance of wakefulness test
PIL, preliminary item list
PSG, polysomnography
ROC, receiver operating characteristic
SART, sustained attention to response task
SDB, sleep disordered breathing
SWAI, sleep-wake activity inventory
SWIFT, sleep-wakefulness inability and fatigue test

REFERENCES

1. Sassani A, Findley LJ, Kryger M, Goldlust E, George C, Davidson TM. Reducing motor-vehicle collisions, costs, and fatalities by treating obstructive sleep apnea syndrome. Sleep 2004;27:453–8.
2. Baldwin CM, Griffith KA, Nieto FJ, O'Connor GT, Walsleben JA, Redline S. The association of sleep-disordered breathing and sleep symptoms with quality of life in the Sleep Heart Health Study. Sleep 2001;24:96–105.
3. Carskadon MA, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S. Guidelines for the Multiple Sleep Latency Test (MSLT): a standard measure of sleepiness. Sleep 1986;9:519–24.
4. Littner MR, Kushida C, Wise M et al., Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. Sleep 2005;28:113–21.
5. Dinges DF, Powell JW. Microcomputer analysis of performance on a portable, simple visual RT task during sustained operations. Behav Res Methods Instrum Comput 1985;6:652–5.
6. Robertson IH, Manly T, Andrade J, Baddeley BT, Yiend J. 'Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia 1997;35:747–58.
7. Sunwoo BY, Jackson N, Maislin G, Gurubhagavatula I, George CF, Pack AI. Reliability of a single objective measure in assessing sleepiness. Sleep 2012;35:149–58.
8. Van Schie MK, Thijs RD, Fronczek R, Middelkoop HA, Lammers GJ, Van Dijk JG. Sustained attention to response task (SART) shows impaired vigilance in a spectrum of disorders of excessive daytime sleepiness. J Sleep Res 2012;21:390–5.
9. Loh S, Lamond N, Dorrian J, Roach G, Dawson D. The validity of psychomotor vigilance tasks of less than 10-minute duration. Behav Res Methods Instrum Comput 2004;36:339–46.
10. Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 1991;14:540–5.

11. Rosenthal L, Roehrs TA, Roth T. The Sleep-Wake Activity Inventory: a self-report measure of daytime sleepiness. Biol Psychiatry 1993;34:810–20.

12. Sangal RB. Evaluating sleepiness-related daytime function by querying wakefulness inability and fatigue: Sleepiness-Wakefulness Inability and Fatigue Test (SWIFT). J Clin Sleep Med 2012;8:701–11.

13. Miletin MS, Hanly PJ. Measurement properties of the Epworth sleepiness scale. Sleep Med 2003;4:195–9.

14. Kendzerska TB, Smith PM, Brignardello-Petersen R, Leung RS, Tomlinson GA. Evaluation of the measurement properties of the Epworth sleepiness scale: a systematic review. Sleep Med Rev 2014;18:321–31.

15. Walter TJ, Foldvary N, Mascha E, Dinner D, Golish J. Comparison of Epworth Sleepiness Scale scores by patients with obstructive sleep apnea and their bed partners. Sleep Med 2002;3:29–32.

16. Sharafkhaneh A, Hirshkowitz M. Contextual factors and perceived self-reported sleepiness: a preliminary report. Sleep Med 2003;4:327–31.

17. Hughes KA. Comparing pretesting methods: cognitive interviews, respondent debriefing, and behavior coding. Statistical Research Division. Washington, DC: US Bureau of the Census, 2004.

18. Nassar-McMillan SC, Borders LD. Use of focus groups in survey item development. Qualitative Report 2002;2:1–11.

19. Cluydts R, De Valck E, Verstraeten E, Theys P. Daytime sleepiness and its evaluation. Sleep Med Rev 2002;6:83–96.

20. Ortiz-Tudela E, Martinez-Nicolas A, Campos M, Rol MA, Madrid JA. A new integrated variable based on thermometry, actimetry and body position (TAP) to evaluate circadian system status in humans. PLoS Comput Biol 2010;6:e1000996.

21. Chiner E, Arriero JM, Signes-Costa J, Marco J, Fuentes I. Validation of the Spanish version of the Epworth Sleepiness Scale in patients with a sleep apnea syndrome. Arch Bronconeumol 1999;35:422–7.

22. Iber C, Ancoli-Israel S, Chesson A, Quan SF, for the American Academy of Sleep Medicine. AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, 1st ed. Westchester, IL: American Academy of Sleep Medicine, 2007.

23. Sangal RB, Thomas L, Mitler MM. Maintenance of wakefulness test and multiple sleep latency test. Measurement of different abilities in patients with sleep disorders. Chest 1992;101:898–902.

24. Lloberes P, Durán-Cantolla J, Martínez-García MA, et al. Diagnosis and treatment of sleep apnea-hypopnea syndrome. Spanish Society of Pulmonology and Thoracic Surgery. Arch Bronconeumol 2011;47:143–56.

25. Weaver TE, Grunstein RR. Adherence to continuous positive airway pressure therapy: the challenge to effective treatment. Proc Am Thorac Soc 2008;15;5:173–8.

26. Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. Trials 2010;11:2.

27. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 2006;4:79.

28. Drake C, Nickel C, Burduvali E, Roth T, Jefferson C, Pietro B. The pediatric daytime sleepiness scale (PDSS): sleep habits and school outcomes in middle-school children. Sleep 2003;26:455–8.

29. Spilsbury JC, Drotar D, Rosen CL, Redline S. The Cleveland adolescent sleepiness questionnaire: a new measure to assess excessive daytime sleepiness in adolescents. J Clin Sleep Med 2007;3:603–12.

30. Kingshott RN, Engleman HM, Deary IJ, Douglas NJ. Does arousal frequency predict daytime function? Eur Respir J 1998;12:1264–70.

31. Wise MS. Objective measures of sleepiness and wakefulness: application to the real world? J Clin Neurophysiol 2006;23:39–49.

32. Sangal RB, Sangal JM, Belisle C. Subjective and objective indices of sleepiness (ESS and MWT) are not equally useful in patients with sleep apnea. Clin Electroencephalogr 1999;30:73–5.

33. Chervin RD, Aldrich MS. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. Neurology 1999;52:125–31.

34. Weaver EM, Kapur V, Yueh B. Polysomnography vs self-reported measures in patients with sleep apnea. Arch Otolaryngol Head Neck Surg 2004;130:453–8.

35. Ballester E, Badia JR, Hernández L et al. Evidence of the effectiveness of continuous positive airway pressure in the treatment of sleep apnea/hypopnea syndrome. Am J Respir Crit Care Med 1999;159:495–501.

36. Engleman HM, Kingshott RN, Wraith PK, Mackay TW, Deary IJ, Douglas NJ. Randomized placebo-controlled crossover trial of continuous positive airway pressure for mild sleep Apnea/Hypopnea syndrome. Am J Respir Crit Care Med 1999;159:461–7.

37. Sangal RB, Thomas L, Mitler MM. Disorders of excessive sleepiness. Treatment improves ability to stay awake but does not reduce sleepiness. Chest 1992;102:699–703.

38. Engleman HM, Martin SE, Deary IJ, Douglas NJ. Effect of continuous positive airway pressure treatment on daytime function in sleep apnoea/hypopnoea syndrome. Lancet 1994;343:572–5.

39. Engleman HM, Cheshire KE, Deary IJ, Douglas NJ. Daytime sleepiness, cognitive performance and mood after continuous positive airway pressure for the sleep apnoea/hypopnoea syndrome. Thorax 1993;48:911–4.

40. Fronczek R, Middelkoop HA, van Dijk JG, Lammers GJ. Focusing on vigilance instead of sleepiness in the assessment of narcolepsy: high sensitivity of the Sustained Attention to Response Task (SART). Sleep 2006;29:187–91.

41. Kushida CA, Nichols DA, Holmes TH, et al. Effects of continuous positive airway pressure on neurocognitive function in obstructive sleep apnea patients: the Apnea Positive Pressure Long-term Efficacy Study (APPLES). Sleep 2012;35:1593–602.

ACKNOWLEDGMENTS

SUBMISSION & CORRESPONDENCE INFORMATION

DISCLOSURE STATEMENT

SUPPLEMENTAL MATERIAL

Table S1—25 prototypical situations related to sleepiness.

- When relaxing
- When standing still
- When eating lunch
- At the table after eating
- When having supper in the evening
- When watching television
- When reading a book, newspaper or magazine
- When writing
- When driving and you are stopped at the lights
- When driving a car
- When riding your motorbike
- When you are in the passenger seat of a car passenger in a car
- When travelling on public transport
- When attending a show or a religious service
- When getting together with family or friends
- When speaking on the phone
- When speaking to someone face to face
- When working
- When in a meeting at work
- When smoking
- When playing cards, dominoes, or on the computer
- When sitting on the toilet
- When washing your face or taking a shower
- When standing, inactive in a public place
- When sitting in a waiting room

Table S2—The preliminary item list format.

	A = Do you feel sleepy or fall asleep?				B = How often does this happen to you?			C = How long does it take before you fall asleep?			Scoring System[†]		
	I do not feel sleepy or fall asleep	I feel sleepy but I do not fall asleep	I feel sleepy and I fall asleep	I fall asleep unexpectedly	Sometimes	Often	Always or most of the time	Less than 5 min	Between 5 and 15 min	More than 15 min	A	A × B	A × B × C
In the morning:													
1. When relaxing	0	1	2	3	1	2	3	3	2	1			
2. When sitting on the toilet	0	1	2	3	1	2	3	3	2	1			
3. When at a boring get-together with family or friends	0	1	2	3	1	2	3	3	2	1			
4. When standing still	0	1	2	3	1	2	3	3	2	1			
5. When standing on public transport	0	1	2	3	1	2	3	3	2	1			
6. When standing and speaking to someone with little interest	0	1	2	3	1	2	3	3	2	1			
7. When driving and you are stopped at the lights	0	1	2	3	1	2	3	3	2	1			
After lunch:													
8. When watching TV with little interest	0	1	2	3	1	2	3	3	2	1			
9. When sitting in a waiting room	0	1	2	3	1	2	3	3	2	1			
10. When driving and you are stopped at the lights	0	1	2	3	1	2	3	3	2	1			
In the afternoon:													
11. When sitting down watching a boring show (theatre, cinema, etc.)	0	1	2	3	1	2	3	3	2	1			
12. When you are in the passenger seat of a car on an interesting trip	0	1	2	3	1	2	3	3	2	1			
13. When sitting down and speaking to someone with little interest	0	1	2	3	1	2	3	3	2	1			
14. When standing, inactive in a public place (waiting to be served, or for the bus, subway, or a friend, etc.)	0	1	2	3	1	2	3	3	2	1			
15. When standing, speaking to someone on the phone with interest	0	1	2	3	1	2	3	3	2	1			
At night:													
16. When having supper	0	1	2	3	1	2	3	3	2	1			
TOTAL SCORE													

[†]Scoring systems: A, severity score; A × B, severity score × frequency of occurrence; A × B × C, severity score × frequency of occurrence × latency to fall asleep. This interviewer-administered questionnaire refers to patient's daily sleepiness evaluated during the "recent weeks." Indicate the appropriate numerical value for the severity of sleepiness referred in each situation **(A)**. Then, indicate the appropriate numerical value for the frequency in which this happens **(B)**. Finally, if the patient falls asleep, indicate how long it takes him to do so **(C)**. If the patient does not feel sleepy or fall asleep in a given situation, score 0 in the first column and continue on to the next situation. If a given situation does not apply to the patient, skip the item completely without scoring.

Table S3—Clinical and PSG changes after CPAP treatment.

	Baseline	CPAP	p
Clinical variables			
Body mass index (kg/m²)	32.2 ± 6.1	32.4 ± 5.6	0.476
HADS-A	5.6 ± 3.4	4.3 ± 3.2	0.022
HADS-D	4.0 ± 3.2	2.6 ± 2.5	0.003
Sleep at home			
Mean nocturnal TST (min)	403.6 ± 49.9	416.9 ± 55.0	0.650
Polysomnographic data			
TIB (min)	472.2 ± 22.6	476.1 ± 18.4	0.496
TST (min)	383.1 ± 48.9	403.6 ± 40.2	0.038
SE (%)	81.0 ± 8.8	84.8 ± 8.2	0.050
WASO (min)	75.4 ± 29.2	64.8 ± 39.5	0.211
Stage 2 sleep latency (min)	15.7 ± 9.2	17.4 ± 23.6	0.692
REM sleep latency (min)	126.2 ± 61.4	98.9 ± 70.9	0.105
Stage 1 (%)	23.9 ± 11.7	13.4 ± 5.6	< 0.001
Stage 2 (%)	55.7 ± 11.6	56.7 ± 8.6	0.680
Stage 3 (%)	6.9 ± 6.3	10.9 ± 7.0	0.007
REM sleep (%)	13.5 ± 5.3	19.0 ± 6.1	0.001
REM episodes (n)	3.3 ± 1.2	3.7 ± 1.1	0.155
Sleep stage changes (n)	212.4 ± 69.5	139.9 ± 39.2	< 0.001
Periodic Limb Movement Index (events/h)	9.6 ± 21.2	21.7 ± 29.6	0.007
Arousal Index (events/h)	48.6 ± 22.2	17.9 ± 5.3	< 0.001
AHI (events/h)	52.5 ± 23.6	2.0 ± 2.2	< 0.001
Mean SpO$_2$ (%)	91.9 ± 2.9	94.9 ± 1.2	< 0.001
CT90 (%)	19.0 ± 20.2	0.3 ± 0.9	< 0.001
Oxygen Desaturation Index 3% (events/h)	43.8 ± 25.9	1.1 ± 1.2	< 0.001
Nadir SpO$_2$ (%)	74.5 ± 9.9	89.5 ± 3.6	< 0.001

PSG, polysomnography; CPAP, continuous positive airway pressure; HADS-A, Hospital Anxiety and Depression Scale - Anxiety; HADS-D, Hospital Anxiety and Depression Scale - Depression; TST, total sleep time; TIB, time in bed; SE, sleep efficiency; WASO, wake after sleep onset; AHI, Apnea-Hypopnea Index; mean SpO$_2$, mean O$_2$ saturation; CT90, cumulative time spend with an O$_2$ saturation below 90%; Nadir SpO$_2$, minimal O$_2$ saturation value; n, number.