ORIGINAL RESEARCH

A pictorial Sleepiness and Sleep Apnoea Scale to recognize individuals with high risk for obstructive sleep apnea syndrome

Cathrin Edelmann[1]
Ramesh Ghiassi[2]
Deborah R Vogt[3]
Martyn R Partridge[2]
Ramin Khatami[4]
Jörg D Leuppi[1,5]
David Miedinger[1,5]

[1]Medical Faculty, University of Basel, Basel, Switzerland; [2]National Heart and Lung Institute, Imperial College London, London, UK; [3]Clinical Trial Unit, University of Basel, University Hospital Basel, Basel, [4]Barmelweid Clinic, Barmelweid, [5]University Clinic of Medicine, Kantonsspital Baselland, Liestal, Switzerland

Correspondence: David Miedinger
University Clinic of Medicine, Kantonsspital Baselland, Rheinstrasse 26, 4410 Liestal, Switzerland
Tel +41 61 925 2525
Email david.miedinger@unibas.ch

Purpose: The aim of this study was to evaluate the validity of a new pictorial form of a screening test for obstructive sleep apnea syndrome (OSAS) – the pictorial Sleepiness and Sleep Apnoea Scale (pSSAS). Validation was performed in a sample of patients admitted to sleep clinics in the UK and Switzerland.

Patients and methods: All study participants were investigated with objective sleep tests such as full-night-attended polysomnography or polygraphy. The pSSAS was validated by taking into account the individual result of the sleep study, sleep-related questionnaires and objective parameters such as body mass index (BMI) or neck circumference. Different scoring schemes of the pSSAS were evaluated, and an internal validation was undertaken.

Results: The full data set consisted of 431 individuals (234 patients from the UK, 197 patients from Switzerland). The pSSAS showed good predictive performance for OSAS with an area under the curve between 0.77 and 0.81 depending on which scoring scheme was used. The subscores of the pSSAS had a moderate-to-strong correlation with widely used screening questionnaires for OSAS or excessive daytime sleepiness as well as with BMI and neck circumference.

Conclusion: The pSSAS can be used to select patients with a high probability of having OSAS. Due to its simple pictorial design with short questions, it might be suitable for screening in populations with low health literacy and in non-native English or German speakers.

Keywords: questionnaire, excessive daytime sleepiness, sleep associated disorders, screening, pictorial, occupation

Introduction

Obstructive sleep apnea syndrome (OSAS) is characterized by repetitive episodes of total and/or partial pharyngeal collapse during sleep.[1] Obesity and age are the best-documented risk factors for this condition.[2] OSAS is estimated to affect 2–15% of middle-aged adults, many of them remaining undiagnosed and untreated.[3–5]

OSAS may disturb sleep architecture leading to fragmented or decreased sleep duration, which results in excessive daytime sleepiness (EDS) and disturbance of concentration. It is associated with high rates of morbidity and mortality mostly due to cardiovascular diseases.[2,6–8] Individuals with OSAS have a six- to sevenfold risk of having a road traffic crash and have an increased risk of cognitive dysfunction, depression and insomnia compared to those without OSAS.[9,10]

The Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine (AASM) has recommended incorporating questions regarding obstructive sleep apnea into routine health evaluations. The diagnostic strategy for OSAS in suspicious patients includes a sleep-oriented history and physical examination and

objective testing.[11] There is no consensus in the American Sleep Association or the AASM about how to screen for OSAS by means of questionnaires.[12]

Mild degrees of OSAS can be treated by lifestyle modifications such as weight loss, abstinence from alcohol/smoking and use of sleep-inducing drugs and tranquilizer.[11,13,14] Continuous positive airway pressure (CPAP) treatment is considered the best treatment option for those with moderate-to-severe OSAS and has been shown to decrease fatal and nonfatal cardiovascular events, normalize driving and attention performances in drivers and reduce overall mortality.[11]

Written questionnaires can be difficult for some patients to complete, and studies have shown that impaired health literacy is common in patients attending a sleep clinic.[15] In addition to the effect of increased sleepiness, verbal episodic memory can be impaired in patients with OSAS.[16] This led to the development of a pictorial Epworth Sleepiness Scale (pESS) in 2011, and the same team further developed a pictorial Sleepiness and Sleep Apnoea Scale (pSSAS) questionnaire that consists of eight items to be used for OSAS screening, four common to the pESS and four related to body morphology, blood pressure and witnessed apnea.[17–21]

The aim of this study was to develop the pSSAS scoring system to discriminate individuals with or without OSAS with a sensitivity and specificity of 80%. For this, we used data of patients with different risk levels for OSAS referred to sleep clinics in the UK and Switzerland.

Methods
The pSSAS questionnaire

The original pSSAS was established for an English-speaking population at the Imperial Sleep Centre at Charing Cross Hospital in London.[18] It is self-administered by the patient and contains eight items (Figure 1A and B). Items 1–4 are used to evaluate EDS and are taken from the pESS.[17] The pictograms show a person in various situations, in particular sitting and reading, watching TV or listening to the radio, sitting quietly after a meal or as a passenger in a car or bus. For each situation, the patient can choose five different integers. Items 5–8 are used to evaluate the risk factors and adverse outcomes of OSAS such as body size, neck size, blood pressure and witnessed pauses while breathing when asleep. Items 1–6 use a Likert scale with five answers, and item 7 allows the individual to choose from four and item 8 from two possible answers. The patient is instructed to select the situation that he/she judges to best represent his/her situation. Details regarding translation and changes in the design of the German version of the pSSAS as well as the suggested scoring scheme can be found in the Supplementary material.

Patient study groups

For this study, the pSSAS was administered between February 2010 and May 2012 to four groups of patients: A) patients examined for OSAS at Charing Cross Hospital, London, UK (Imperial College Healthcare NHS Trust); B) patients examined for OSAS at the sleep clinic of the Barmelweid Clinic, Barmelweid, Switzerland; C) patients admitted to the emergency department at the University Hospital Basel, Basel, Switzerland, for treatment and investigation after an occupational injury; and D) outpatients referred to the surgical outpatient or the occupational medicine department at the University Hospital Basel, Basel, Switzerland, for reasons other than injuries due to occupational accidents.

Questionnaires

For all patients, age, gender, height and weight were recorded. Each patient completed the English or German pSSAS as well as the English or German version of the ESS.[22] Patients of groups B–D also answered the German version of the Sleep Apnea Scale of the Sleep Disorders Questionnaire (SA-SDQ).[23] A total of n=26 individuals from patient groups A and B were asked to fill in the pSSAS a second time (retest) before a treatment was initiated.

Scoring of sleep studies and attribution of OSAS diagnosis

The diagnosis of OSAS was based on an overnight sleep study. Information regarding scoring of sleep studies can be found in the Supplementary material. Regarding the presence of OSAS, the following criteria were considered in this study: No OSAS if apnea–hypopnea index (AHI) <5 or AHI 5–14 and ESS score <10; OSAS if AHI ≥15 or AHI 5–14 and ESS score ≥10.

pSSAS scoring scheme

Table 1 shows the three different scoring schemes which were evaluated. They differ in the scoring of pSSAS items 7 and 8.

Statistical analysis

Only patients with no missing values in the pSSAS were included in the full analysis set. Descriptive statistics of demographic and baseline characteristics were calculated for the four patient groups separately and for the full analysis set. The performance of the different pSSAS scoring schemes in predicting OSAS was investigated by fitting separate single-predictor logistic regression models. For each pSSAS scoring scheme, the area under the receiver operating characteristic (ROC) curve (AUC) was determined. Since the three ROC
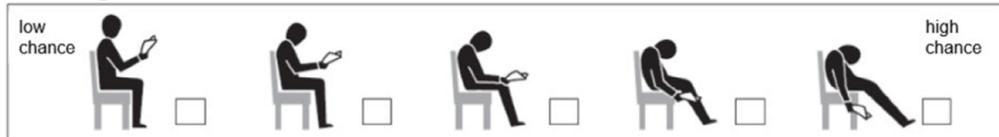
A Patient name_____ Hospital Number_____ Date of Birth _____ Today's Date_____

The pictorial Sleepiness and Sleep Apnoea Scale (pSSAS)

Tick ☑ **one box only** in **each different situation** that best represents you.

How likely are you to <u>doze off or fall asleep</u> in the following situations?
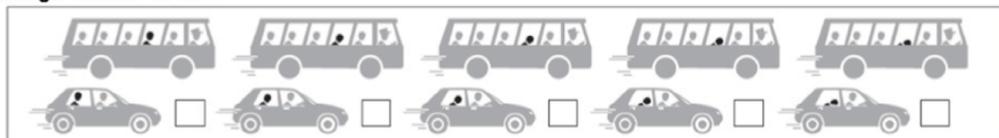
1. Sitting and Reading

2. Watching TV or listening to the radio

3. Sitting quietly after a meal
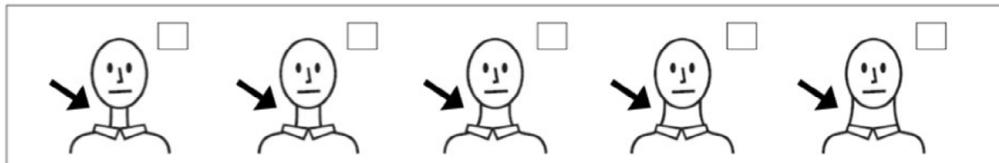

4. As a passenger in a car or bus

Now please tick ☑ **one box only** in **each series of pictures** that best represents you.

5. Body size
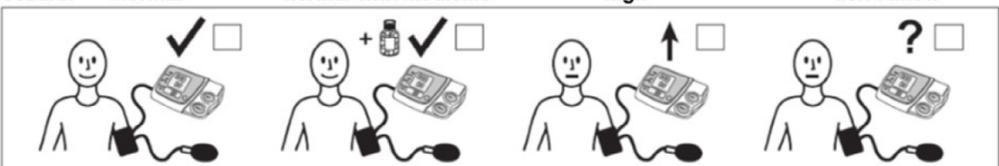

6. Neck size
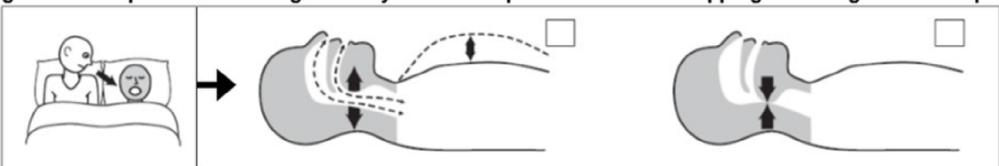

7. Blood pressure: normal   normal with medicine   high   don't know

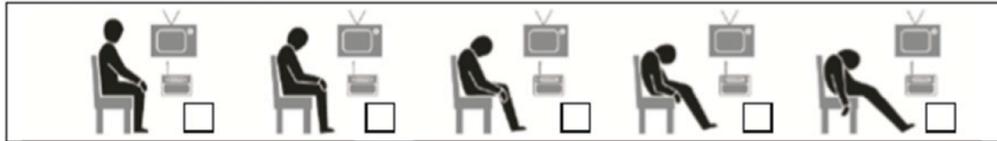### 8. Breathing when asleep: breathing normally when asleep   stopping breathing when asleep

©

Figure 1 (Continued)

B Markieren sie bitte in jeder Situation <u>ein</u> Kästchen mit ☑, das ihnen am ehesten entspricht.
Wie leicht fällt es ihnen in folgenden Situationen <u>einzuschlafen</u>?
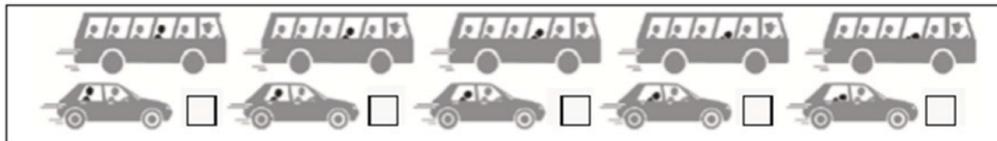
1. Lesen und Sitzen

2. Fernsehen oder Radio hören

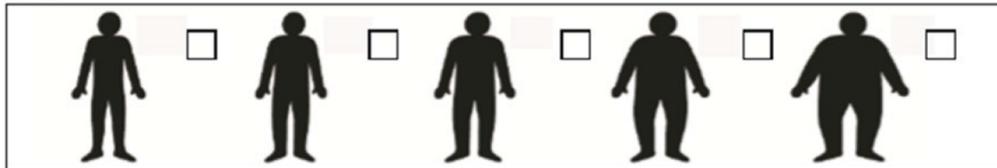3. Ruhiges Sitzen nach dem Essen

4. Als Mitfahrer in einem Auto oder Bus
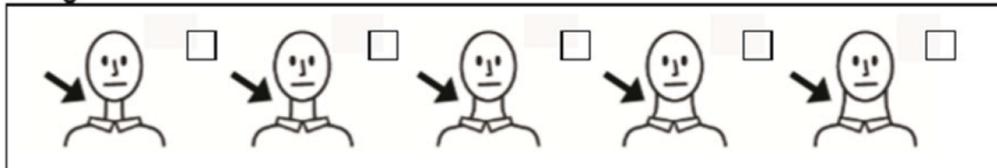
Markieren sie bitte hier <u>ein</u> Kästchen mit ☑, das ihnen am ehesten entspricht.
5. Körperfülle

6. Halsumfang

7. Blutdruck: normal   normal mit Medikamenten   hoch   weiss es nicht
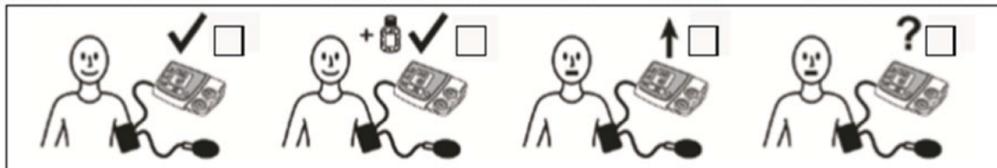
8. Atmung im Schlaf   normale Atmung im Schlaf   Atmungspausen im Schlaf   weiss es nicht
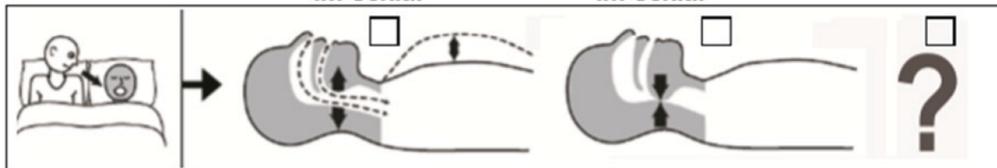
Figure 1 (**A**) English and (**B**) German version of the pictorial Sleepiness and Sleep Apnoea Scale questionnaire.
Notes: The English version of the pSSAS questionnaire provided with permission by Dr Ghiassi. The German version of the pSSAS questionnaire provided with permission by Dr Miedinger and Dr Edelmann.

Table 1 pSSAS scoring schemes

pSSAS item		Scoring
1–4	Daytime sleepiness	0, 1, 2, 3 and 4 points with increasing score indicating increasing daytime sleepiness
5	Body size	0, 1, 2, 3 and 4 points with increasing score indicating increasing body size
6	Neck size	0, 1, 2, 3 and 4 points with increasing score indicating increasing neck size
7	Blood pressure	Normal =0 point Normal with antihypertension medication or high blood pressure =4 points Do not know =0 point (for scoring scheme 1), 2 points (for scoring scheme 2) and 4 points (for scoring scheme 3)
8	Sleep apnea	Normal breathing =0 point Breathing pauses do occur =4 points Do not know =0 point (for scoring scheme 1), 2 points (for scoring scheme 2) and 4 points (for scoring scheme 3)

Abbreviation: pSSAS, pictorial Sleepiness and Sleep Apnoea Scale.

curves cross within sensitivity 80–100%, the corrected partial AUCs for sensitivity 80–100% are given as well. For each ROC curve, the "optimal" cutoff value, that is, the score yielding the highest combination of sensitivity and specificity, and the cutoff value with the highest specificity for a given sensitivity of at least 80% ("80% sensitivity") were determined. The 95% confidence intervals of AUCs, sensitivities and specificities were determined using a bootstrap resampling method (k=10000). To get a more accurate estimate of the model performance, an internal validation was performed. An ordinary bootstrap resampling procedure (k=1000) following

the recommendations of Steyerberg et al was applied.[24] Each pSSAS scoring scheme was validated on the full data set and on all four patient groups. As a measure of discrimination, bias-corrected Somers' D_{xy} rank correlation is presented. D_{xy} corresponds to 2*(AUC–0.5) and thus ranges from 0 (random predictions) to 1 (perfectly discriminating predictions). As a measure of calibration, we present the calibration slope. This is the corrected slope of the overall logistic calibration equation observation$_i$ ~ prediction$_i$. It can be thought of as a shrinkage factor, accounting for overfitting. Well-calibrated models have a slope of 1. Finally, as a measure of the overall accuracy, we present the Brier score. It represents the average prediction error and ranges from 0 (perfect model) to 0.25 (useless model). Test–retest reliability of the pSSAS after a certain time without initiating treatment for OSAS was estimated using either Spearman's rank correlation coefficient (items 1–6 and total pSSAS score) or the proportion of identical scores (items 7 and 8). All statistical analyses were performed using R (R Core Team, 2014, version 3.0.1). The R package pROC was used for all analyses and figures of ROC curves.[25] Internal validation was done using the R package rms (R Core Team, 2013, package version 3.6.3).

Results
Demographic and baseline characteristics

Eleven patients with one or more missing answers were excluded (group A =0, group B =5, group C =5, group D =1). The full data set consisted of 431 individuals: patient group A=234 (54.3%), group B=82 (19%), group C=79 (18.3%), and group D=36 (8.4%) individuals. Clinical data, questionnaire and results of sleep studies can be seen in Table 2. Prevalence of OSAS and mean body mass index (BMI) was highest in the groups A and B. The groups A–C consisted of mainly male individuals.

Table 2 Demographic and baseline characteristics according to patient group

	Group A	Group B	Group C	Group D	All	*p*-value
Male sex	163 (70%)	55 (67%)	44 (56%)	18 (50%)	280 (65%)	p=0.03
OSAS diagnosis	171 (73.1%)	47 (57.3%)	11 (13.9%)	4 (11.3%)	233 (54.1%)	p=0.0005
Age (years)	46.6 (17.8)	51.9 (16.0)	36.7 (21.0)	43.2 (18.5)	45.5 (21.0)	$p<0.0001$
BMI (kg/m²)	33.6 (10.1)	30.7 (7.7)	26.3 (6.0)	25.1 (4.7)	31 (8.9)	$p<0.0001$
AHI (events/hour)	26.5 (31)	18.2 (20.8)	6.9 (6.0)	3.2 (4.0)	19.4 (24.5)	$p<0.0001$
ODI (events/hour)	20.8 (24.0)	24.5 (23.0)	4.3 (4.0)	2.7 (3.0)	17.0 (21.0)	$p<0.0001$
ESS (points)	10.9 (8.0)	10.0 (8.0)	7.3 (4.0)	6.9 (4.0)	9.7 (7.0)	$p<0.0001$
SA-SDQ (points)	NA	41.8 (12.0)	29.3 (9.0)	27.7 (10.5)	33.9 (13.0)	$p<0.0001$

Notes: Group A: Patients examined for OSAS at Charing Cross Hospital, London; Group B: patients examined for OSAS at the Sleep Clinic of the Barmelweid Clinic, Switzerland; Group C: patients with injuries admitted to the Emergency Department at the University Hospital Basel, Switzerland; and Group D: outpatients referred to the Surgical Outpatient or the Occupational Medicine Department at the University Hospital Basel, Switzerland. Data are presented as frequency (%) or median (interquartile range). p-values for equivalence of all groups were calculated using Fisher's exact test or Kruskal–Wallis test.
Abbreviations: OSAS, obstructive sleep apnea syndrome; BMI, body mass index; AHI, apnea–hypopnea index; ODI, oxygen desaturation index; ESS, Epworth Sleepiness Scale; SA-SDQ, Sleep Apnea Scale of the Sleep Disorders Questionnaire; NA, not available.

Development and validation

The frequency distributions of the three pSSAS scoring schemes according to OSAS diagnoses can be seen in Figure 2. For each scoring scheme, there was a considerable overlap between positive and negative OSAS diagnoses. The predictive performance of the three different pSSAS scoring schemes is shown in Table 3. Overall, the diagnostic performance of pSSAS scoring scheme 1 was slightly better than the scoring schemes 2 and 3 (Table 4 and Figure 3A and B). The sensitivity is between 0.670 and 0.839 and specificity is between 0.579 and 0.812 depending on the chosen cutoff (between 12 and 14 points).

However, none of the three pSSAS scoring schemes achieved the aim of discriminating negative and positive OSAS diagnoses with a sensitivity and specificity of 80% each. As sensitivity was given more weight than specificity, an additional cutoff value was determined – 80 se cutoff value – which corresponds to the cutoff value with the highest specificity for a given sensitivity of at least 80% (Table 3). This is in accordance with the ROC analyses described in the previous section (Figure 3A and B). The discriminatory properties, that is, the ability to distinguish high-risk individuals from low-risk individuals, of the three pSSAS scoring schemes varied between the data sets. They were best in the full analysis set (groups A–D) with pSSAS scoring scheme 1 (D_{xy}=0.61, which corresponds to an AUC of 0.81; Table 3). Discrimination using the scoring scheme 1 was worst in patient group C (D_{xy}=0.34). The calibration of all pSSAS
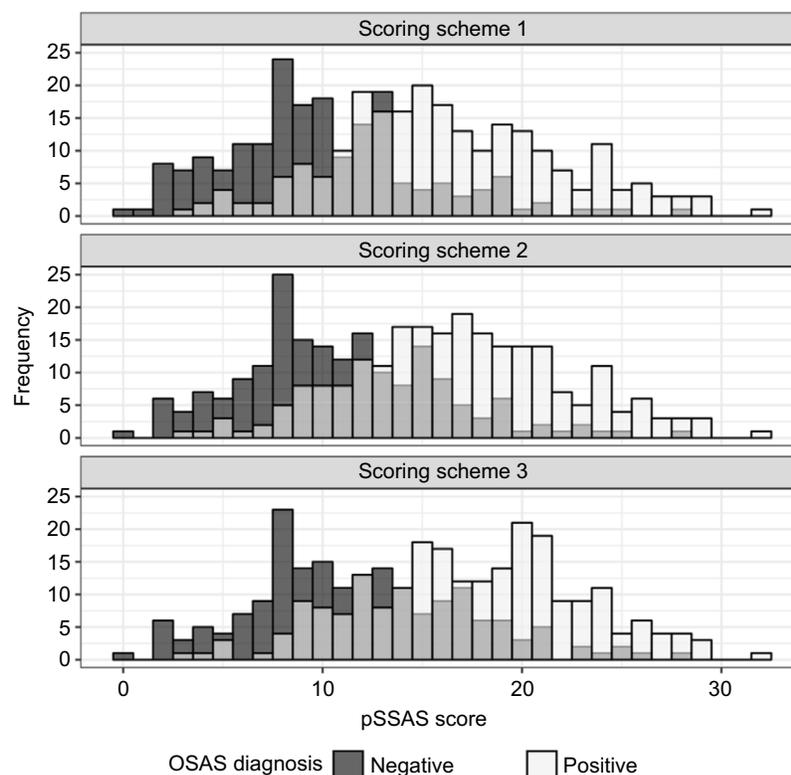

Figure 2 Frequency distribution of the pSSAS questionnaire scores for individuals according to OSAS diagnosis.
Notes: White = individuals with OSAS; black = individuals without OSAS; gray = overlap of the columns.
Abbreviations: pSSAS, pictorial Sleepiness and Sleep Apnoea Scale; OSAS, obstructive sleep apnea syndrome.

Table 3 Predictive performance of the three different pSSAS scoring schemes

	AUC	Cutoff	Value	Sensitivity	Specificity
Score 1	0.805 (0.763–0.847)	Optimal	14	0.670 (0.609–0.730)	0.821 (0.768–0.874)
		80 sensitivity	12	0.822 (0.770–0.870)	0.647 (0.579–0.711)
Score 2	0.794 (0.752–0.837)	Optimal	12	0.839 (0.791–0.887)	0.579 (0.505–0.647)
		80 sensitivity	12	0.839 (0.791–0.887)	0.579 (0.505–0.647)
Score 3	0.774 (0.730–0.819)	Optimal	14	0.761 (0.704–0.813)	0.658 (0.589–0.726)
		80 sensitivity	13	0.796 (0.743–0.848)	0.584 (0.516–0.653)

Notes: AUC, sensitivity and specificity of the best cutoff value and lowest cutoff value with a minimal sensitivity of 80% are given. 95% bootstrap confidence intervals are given in parentheses.
Abbreviations: pSSAS, pictorial Sleepiness and Sleep Apnoea Scale; AUC, area under the receiver operating characteristic curve.

Table 4 Internally validated performance of all three pSSAS scoring schemes for the full data set and for each patient group

Data set	pSSAS scoring scheme	D_{xy}	Calibration slope	Brier score
Full	1	0.61	1.01	0.18
	2	0.59	1.00	0.19
	3	0.55	1.01	0.19
A	1	0.49	1.01	0.17
	2	0.50	1.01	0.17
	3	0.50	1.02	0.17
B	1	0.50	1.04	0.21
	2	0.51	1.07	0.20
	3	0.49	1.07	0.21
C	1	0.34	1.05	0.12
	2	0.38	1.26	0.12
	3	0.39	1.24	0.12
D	1	0.40	1.05	0.11
	2	0.56	1.50	0.11
	3	0.60	1.32	0.10

Notes: D_{xy} =bias-corrected Somers' D_{xy} rank correlation. Group A: Patients examined for OSAS at Charing Cross Hospital, London; Group B: patients examined for OSAS at the Sleep Clinic of the Barmelweid Clinic, Switzerland; Group C: patients with injuries admitted to the Emergency Department at the University Hospital Basel, Switzerland; and Group D: outpatients referred to the Surgical Outpatient or the Occupational Medicine Department at the University Hospital Basel, Switzerland.
Abbreviation: pSSAS, pictorial Sleepiness and Sleep Apnoea Scale.

scoring schemes was very good in the full analysis set and in patient groups A and B (Table 4; calibration slope closely around 1). The worst calibration was found for pSSAS scoring scheme 2 applied to patient group D (slope=1.5). These differences can be explained by the differences in the size of the data sets. The pSSAS scoring scheme 1 performed rather well in patient groups C and D (slope=1.05). Brier score differed notably between the data sets and was quite high for patient group B (B around 0.2; Table 4). Lowest Brier scores were observed in patient group D (B around 0.1).

Internal consistency was evaluated by pairwise correlation of pSSAS items (Figure 4). The pSSAS items 1–4 evaluate EDS. While a rather strong correlation was found for pSSAS items 1 ("likeliness of falling asleep while sitting/reading") and 3 ("sitting quietly after having eaten"; Spearman's rank correlation coefficient $\rho=0.6$), the other items correlations were only weak or moderate. The pSSAS items 5 (body size) and 6 (neck circumference) correlated strongly ($\rho=0.7$). Construct validity was investigated by correlating pSSAS items or scores with other established measures for EDS or OSAS risk: the pSSAS subscore "excessive daytime sleepiness" correlated strongly with ESS total score ($\rho=0.8$; Figure 5). Items 5 (body size) and 6 (neck circumference) correlated strongly with BMI ($\rho=0.8$) and measured neck circumference ($\rho=0.5$), respectively. With data from patient groups B–D, the correlation of the pSSAS subscore of items 5–8 "risk factors and adverse outcomes" with total SA-

SDQ score was determined. The different scoring schemes resulted in a correlation score of $\rho=0.5$ for all scoring schemes (1, 2 and 3) with SA-SDQ total score (Figure 6).

Twenty-six individuals completed the pSSAS for a second time prior to undergoing any treatment for OSAS. The median number of days between the tests was 13.5 days (range 0–90 days). Items 1–6 showed quite strong correlations between test and retest scores, ranging from $\rho=0.6$ (item 1) to $\rho=0.9$ (item 4). For item 7 (blood pressure), the consistency between test and retest (percentage of identical answers) was 58%. For item 8 (witnessed apnea during sleep), the consistency between test and retest was 73%. Test–retest correlations of pSSAS subscore "excessive daytime sleepiness", pSSAS subscore "risk factors and adverse outcomes" and the total pSSAS score were $\rho=0.8$, $\rho=0.7$ and $\rho=0.7$ for the pSSAS scoring schemes (1, 2 and 3), respectively (Figure 7).

Discussion

A tool for quantification of EDS is essential in sleep medicine, and the ESS is widely used. However, self-completion may be problematical for those with impaired health literacy, and a pictorial translation of the Epworth score has been demonstrated to produce comparable results.[18,20] The pSSAS showed good predictive performance with an AUC between 0.77 and 0.81 depending on the scoring scheme used. The subscores of the pSSAS had a moderate-to-strong correlation with widely used screening questionnaires for OSAS or EDS like the SA-SDQ or the ESS, respectively. The pSSAS showed good retest reliability for most of the individual items, the subscores and the total score.

Given the limited availability of full-night-attended sleep studies, an easy-to-administer instrument that allows quantification of EDS and identification of individuals with a high risk for OSAS would be a useful tool in clinical medicine, and several written questionnaires or tools exist, some to be self-completed and others physician administered.[25,26] However, when using tools such as questionnaires, we must acknowledge that patients may struggle to understand the meaning of the questions or be unable to read. Ghiassi and Partridge have shown that 16% of patients attending a London sleep center showed impaired health literacy when investigated with the Rapid Estimate of Adult Literacy in Medicine (REALM) tool.[15] This finding was surprising as all the patients admitted using English as their everyday language. The authors offer two explanations for the low REALM scores: the fact that a significant proportion of these patients were non-native English speakers or that cognitive defects associated with undiagnosed OSAS and EDS could have contributed to this result. An impairment in logical memory test results was pre-
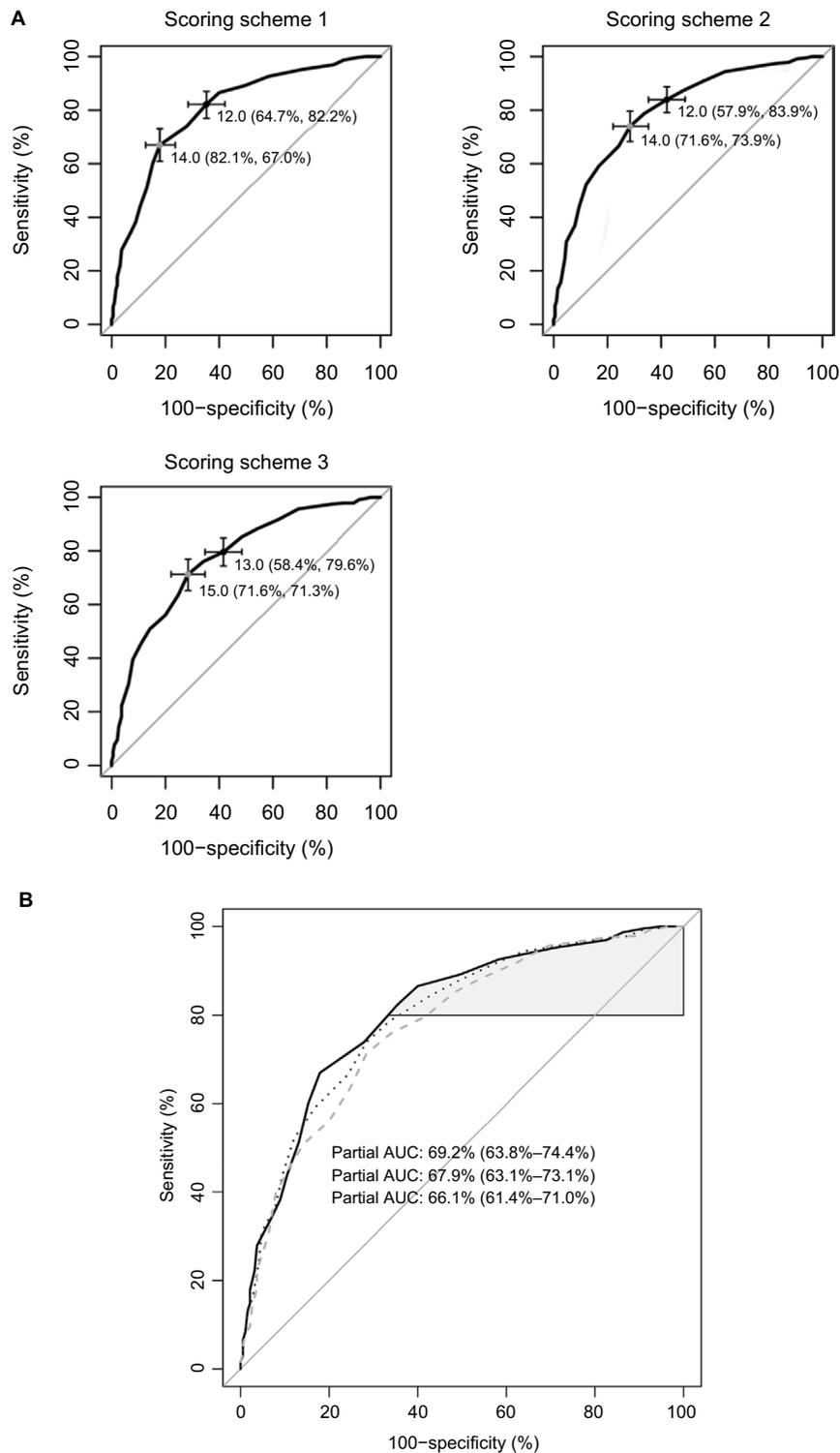
Figure 3 (**A**) Receiver operating characteristic curves for the three different pSSAS scoring schemes. (**B**) Corrected partial AUC for 80–100% sensitivity for pSSAS scoring schemes 1 (black, continuous line), 2 (dark gray, dotted line) and 3 (light gray, dashed line).
Notes: (**A**) Cutoff values ("optimal" in gray; "80% sensitivity" in black) are shown with 95% confidence intervals for sensitivity and 1−specificity. Numbers indicate the pSSAS score with specificity and sensitivity in brackets. (**B**) Numbers in brackets indicate the 95% confidence interval of the partial AUC.
Abbreviations: pSSAS, pictorial Sleepiness and Sleep Apnoea Scale; AUC, area under the receiver operating characteristic curve.

viously shown in patients with OSAS.[16] It has been estimated that in primary schools in England, one in nine children do not speak English as the first language.[27] Therefore, a short and easy-to-complete questionnaire is needed. Ghiassi et al have shown that a pictorial version of the Epworth Sleepiness Scale was rated very easy to complete and preferred
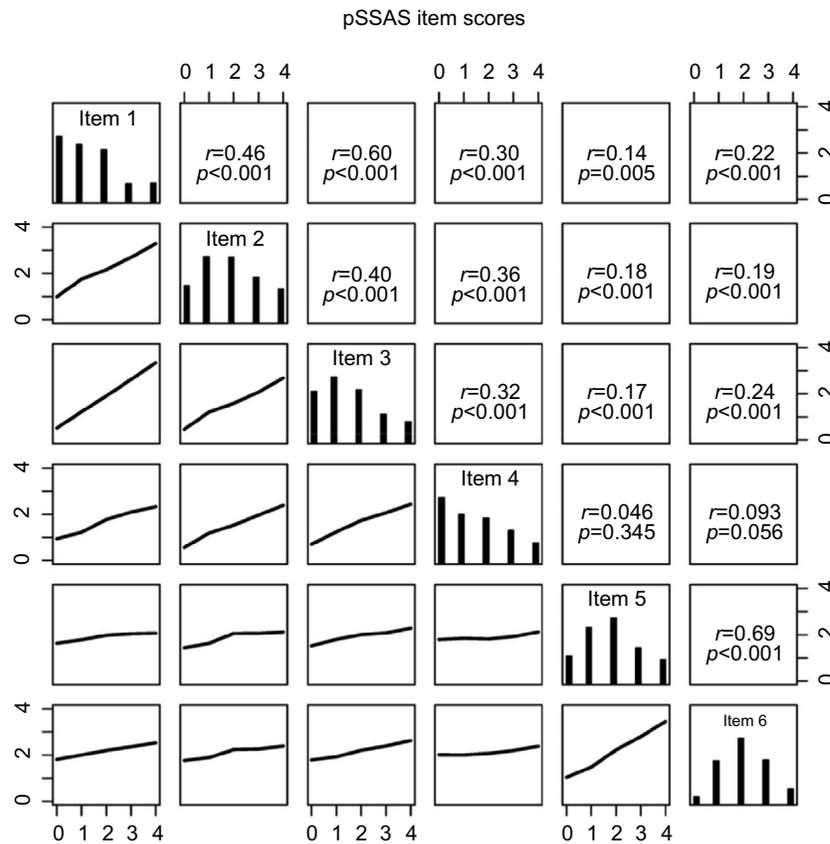
Figure 4 Pairwise correlation of pSSAS items 1–6.
Notes: The diagonal shows the frequency distribution of the scores of each item. Plots below the diagonal show a smoother of the pairwise linear relationships with the axes indicating the score points (0–4) of the items.
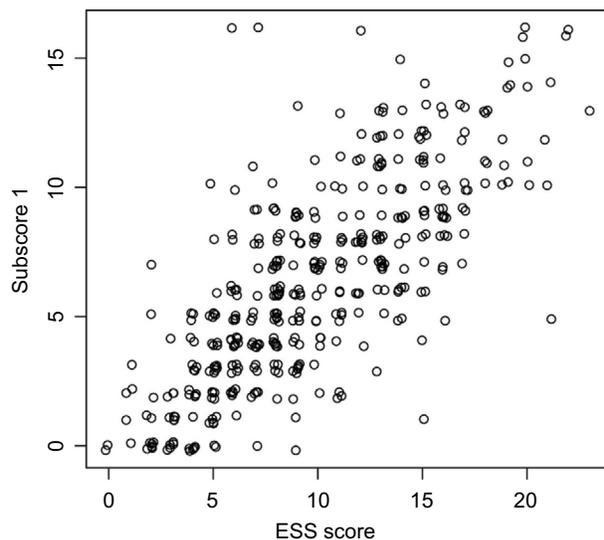**Abbreviation:** pSSAS, pictorial Sleepiness and Sleep Apnoea Scale.

Figure 5 Correlation of pSSAS "excessive daytime sleepiness" subscore with ESS score.
Notes: n=411, ρ=0.76, p<0.001. To prevent overplotting, data points are slightly jittered.
Abbreviations: pSSAS, pictorial Sleepiness and Sleep Apnoea Scale; ESS, Epworth Sleepiness Scale.

Figure 6 Correlation of pSSAS "risk factors and adverse events" subscore with SA-SDQ score according to pSSAS scoring schemes 1, 2 and 3.
Notes: n=117, ρ=0.53, ρ=0.54 and ρ=0.50, p<0.001 for all correlations. To prevent overplotting, data points are slightly jittered.
Abbreviations: pSSAS, pictorial Sleepiness and Sleep Apnoea Scale; SA-SDQ, Sleep Apnea Scale of the Sleep Disorders Questionnaire.

to the original version by respondents.[17] However, even though the ESS is often used to screen for OSAS, it was not developed for this purpose and has its limitation. Bloch et
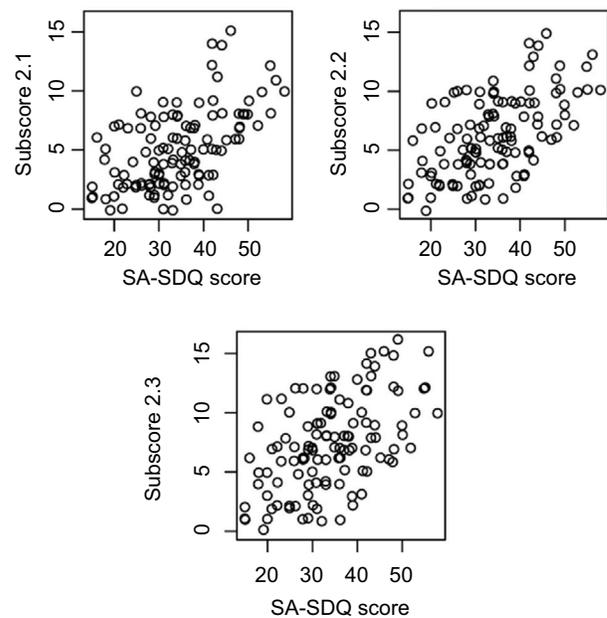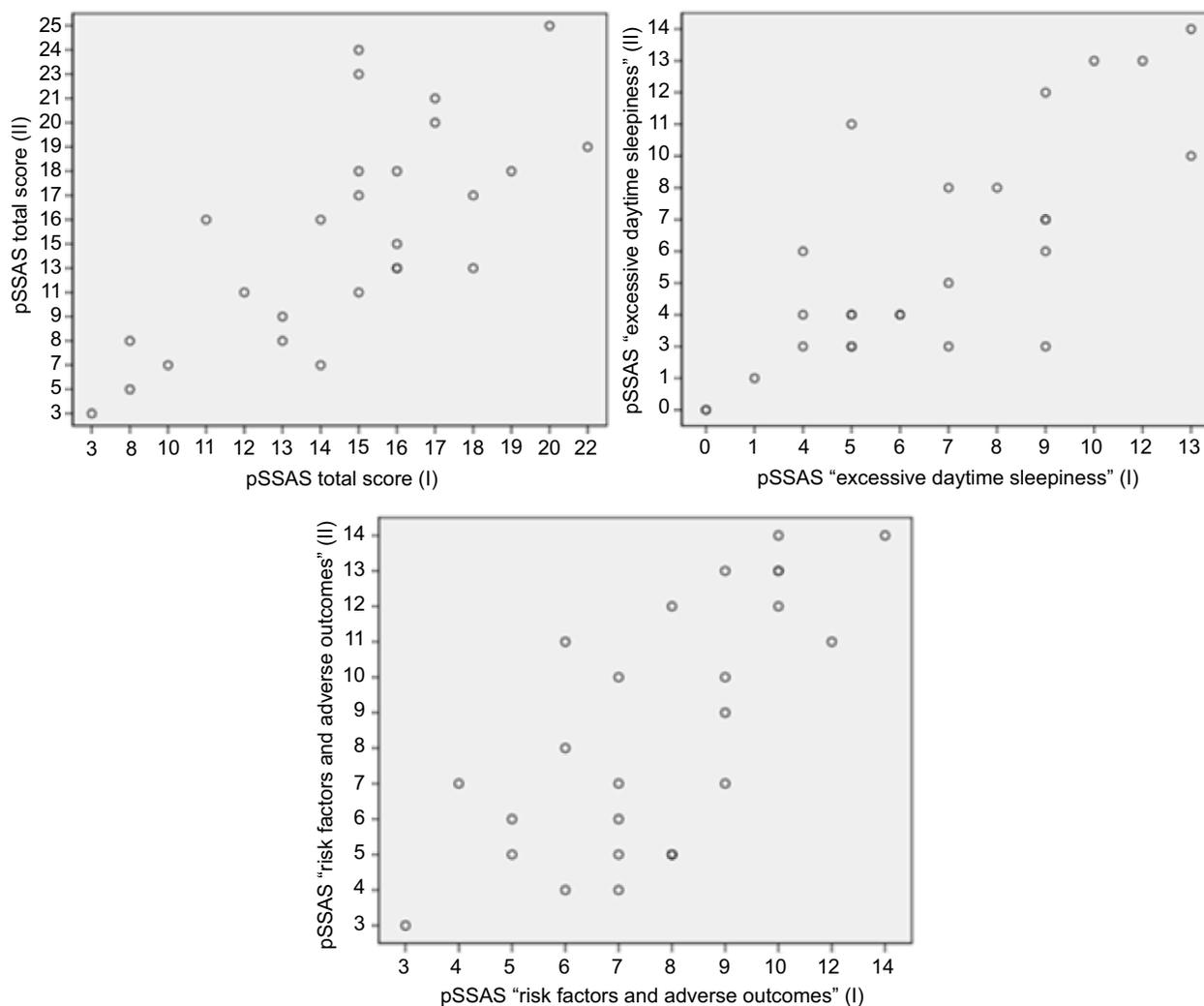
Figure 7 Test–retest correlation of pSSAS subscores and total score.
Note: (I)=questionnaire answered on the first day; (II)=questionnaire answered on a consecutive day prior to treatment for OSAS.
Abbreviations: pSSAS, pictorial Sleepiness and Sleep Apnoea Scale; OSAS, obstructive sleep apnea syndrome.

al evaluated the ESS in a German population in 1999 and found a relatively weak correlation of the ESS scores with the respiratory disturbance index.[22] This suggests that the ESS needs to be combined with other predictors or measures to be more useful in screening populations for OSAS. Other questionnaires such as the SA-SDQ are more extensive and complex and include items on body weight, age, sex, smoking status as well as sleep disturbances. However, due to the SA-SDQ's complexity and length, not all items are always completed especially if the questionnaire is filled in by the patient itself without supervision and the possibility to clarify questions. In the original validation study of this questionnaire by Douglass et al, only 94% of items were completed by the respondents.[23]

The three pSSAS scoring schemes performed equally on the full data set and on the four patient groups separately, and hence, none of the three scores can be declared as best. Performance differed between the four patient groups, but there was no distinct pattern when looking at all three measures D_{xy}, slope and B. Discrimination (D_{xy}) and calibration was best in patient group A, while overall accuracy B was best in patient group D. The performance on the full data set can be summarized as fair to good. The diagnostic performance of pSSAS scoring scheme 1 was slightly better than the scoring schemes 2 and 3. None of the three pSSAS scoring schemes proved to be superior. However, we suggest using the scoring scheme 1 given its superior performance in the overall sample investigated in our study. We suggest using a cutoff of 14 points with a sensitivity of almost 70% and a specificity of >80%. A cutoff of 12 points in our population increased sensitivity to >80% but decreased specificity to 64%. Despite its frequent use to screen patients for OSAS, the conventional ESS has been shown to have low sensitivity and moderate specificity (sensitivity 38–53% and specificity 59–78%) in individuals having EDS and being investigated with polysomnography or polygraphy.[28,29]

The pSSAS sleepiness component correlated strongly with the ESS. This is in accordance with the original validation study of the pESS where an acceptable level of equivalence and comparable total ESS scores were obtained between the two questionnaires.[17] pSSAS items 5 and 6 require self-assessment of body size and neck size. There was a high correlation of body size with BMI but only moderate correlation of neck size with measured neck circumference. It has been shown that misclassification of weight and obesity status can occur if adolescents are asked to self-report height, weight and weight perception.[30] However, it is unclear why the correlation between self-assessment of neck size and measured neck size is worse as compared to body size.

In a subset of patients from Switzerland, we performed test–retest analysis of pSSAS item, subscore and total score. Patients filled in the questionnaire on the first day of the evaluation for OSAS before any measurements were made. However, some of these individuals were aware of blood pressure and even sleep study results the second time the questionnaire was answered. Therefore, we cannot exclude that this new information led to a different answering pattern when the pSSAS was filled in the second time. The correlation of responses was moderate or even high for the items forming the sleepiness component and the items on body and neck size. Test–retest was previously performed in a sample of patients investigated in the UK. The correlation coefficient for the sleepiness and risk components as well as the total score was similar to the Swiss patients.[18]

For this study, we considered OSAS to be clinically relevant in individuals with AHI 5–15 events/hour and significantly increased daytime sleepiness equivalent to an ESS score of ≥10. We considered an individual to have clinically significant OSAS if during the sleep study an AHI ≥15 was determined. Individuals with lower AHI (AHI ≥5) but with clinically important daytime sleepiness (ESS ≥10) were also considered to have OSAS as these patients may be considered candidates for interventions such as lifestyle modifications, weight reduction, positional therapy or maybe even a CPAP treatment trial.[31,32] However, there is ongoing debate about the management of patients with mild OSAS.[33,34] In our opinion, the decision to investigate further with objective sleep studies or to start treatment should always be based on individual risk factors and patient preferences.

Trained board-certified sleep specialists or pneumologists according to relevant scoring rules without knowledge of the individual pSSAS scores rendering differential bias unlikely performed scoring of the sleep studies. However, due to the different setting in the three sleep centers in terms of organiza-tion, devices used for the sleep studies and the fact that sleep studies were scored by different persons, observer bias cannot be fully excluded. The original form of the pSSAS used in UK patients did not allow individuals to choose "don't know" when asked about apneas during sleep. This answering option was added to the German version of the pSSAS. Therefore, different scoring schemes were only validated in a subset of the patients. To assign a diagnosis of OSAS, we considered EDS measured with the ESS and the results of the full-night-attended polysomnography. The ESS is regularly used to measure and rate the severity of EDS in all three study centers. We decided to use a transparent and widely recognized definition for OSAS that we could apply to all data sets regardless of the health system or practice of care. However, the first part of the pSSAS (items 1–4) is very similar to the conventional ESS where patients need to rate the probability to fall asleep in standard situations. The situation that parts of the test under study are partially included in the reference standard could lead to an overestimation of the pSSAS diagnostic accuracy.

As a next step, one could evaluate if different cutoff scores according to sex or age of the patient increase the diagnostic properties of the pSSAS, and if the results of validity and the use of the pSSAS in individuals with measured low health literacy are as good as in our unselected population. Further, one should investigate if the pSSAS score is responsive to treatment interventions, and it should be made available for other languages and tested in other clinical settings.

The pSSAS is a novel tool that can be used to screen individuals with EDS in an outpatient clinic in English- and German-speaking setting. Its simple structure, scoring rules and the use of pictograms make it an alternative to more complex and long questionnaires in clinical practice and in patients with impaired literacy. The validation study has shown good internal consistency, repeatability and construct validity when compared with established tools for the diagnosis of EDS and OSAS. However, further research is needed to evaluate its responsiveness to therapy and its use in other settings such as population screening or self-administration without supervision.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Badr MS. Pathophysiology of obstructive sleep apnea. *Oral Maxillofac Surg Clin North Am*. 2002;14(3):285–292.
2. Yaggi HK, Concato J, Kernan WN, Lichtman JH, Brass LM, Mohsenin V. Obstructive sleep apnea as a risk factor for stroke and death. *N Engl J Med*. 2005;353(19):2034–2041.

3. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med*. 1993;328(17):1230–1235.

4. Heinzer R, Vat S, Marques-Vidal P, et al. Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. *Lancet Respir Med*. 2015;3(4):310–318.

5. Gibson GJ. Obstructive sleep apnoea syndrome: underestimated and undertreated. *Br Med Bull*. 2005;72:49–65.

6. Redline S, Yenokyan G, Gottlieb DJ, et al. Obstructive sleep apnea-hypopnea and incident stroke: the sleep heart health study. *Am J Respir Crit Care Med*. 2010;182(2):269–277.

7. Lavie P, Herer P, Peled R, et al. Mortality in sleep apnea patients: a multivariate analysis of risk factors. *Sleep*. 1995;18(3):149–157.

8. Bassetti CL, Milanova M, Gugger M. Sleep-disordered breathing and acute ischemic stroke: diagnosis, risk factors, treatment, evolution, and long-term clinical outcome. *Stroke*. 2006;37(4):967–972.

9. George CF. Reduction in motor vehicle collisions following treatment of sleep apnoea with nasal CPAP. *Thorax*. 2001;56(7):508–512.

10. Chokroverty S. Overview of sleep & sleep disorders. *Indian J Med Res*. 2010;131:126–140.

11. Epstein LJ, Kristo D, Strollo PJ Jr, et al; Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med*. 2009;5(3):263–276.

12. Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. *Anesthesiology*. 2009;110(4):928–939.

13. Randerath WJ, Verbraecken J, Andreas S, et al; European Respiratory Society task force on non-CPAP therapies in sleep apnoea. Non-CPAP therapies in obstructive sleep apnoea. *Eur Respir J*. 2011;37(5):1000–1028.

14. Kline CE, Crowley EP, Ewing GB, et al. The effect of exercise training on obstructive sleep apnea and sleep quality: a randomized controlled trial. *Sleep*. 2011;34(12):1631–1640.

15. Ghiassi R, Partridge MR. Health literacy and sleep apnoea. *Thorax*. 2011;66(2):180.

16. Twigg GL, Papaioannou I, Jackson M, et al. Obstructive sleep apnea syndrome is associated with deficits in verbal but not visual memory. *Am J Respir Crit Care Med*. 2010;182(1):98–103.

17. Ghiassi R, Murphy K, Cummin AR, Partridge MR. Developing a pictorial Epworth Sleepiness Scale. *Thorax*. 2011;66(2):97–100.

18. Ghiassi R. *The Development of Pictorial Tools for Obstructive Sleep Apnoea Syndrome*. London: Imperial College London; 2014.

19. Drakatos P, Ghiassi R, Jarrold I, et al. The use of an online pictorial Epworth Sleepiness Scale in the assessment of age and gender specific differences in excessive daytime sleepiness. *J Thorac Dis*. 2015;7(5):897–902.

20. Fairley D, Vennelle M, Azuaje-Borges R, Riha R. Comparison of the pictorial Epworth sleepiness scale and the Epworth sleepiness scale in a sleep clinic population. *Eur Respir J*. 2013;42 Suppl 57:P2030.

21. Karim A, Arora VK, Gupta MB. Emerging applications: screening OSA by Modified Pictorial Epworth Sleepiness Scale in Indian subjects. *Indian J Tuberc*. 2015;62(4):222–225.

22. Bloch KE, Schoch OD, Zhang JN, Russi EW. German version of the Epworth Sleepiness Scale. *Respiration*. 1999;66(5):440–447.

23. Douglass AB, Bornstein R, Nino-Murcia G, et al. The Sleep Disorders Questionnaire. I: creation and multivariate structure of SDQ. *Sleep*. 1994;17(2):160–167.

24. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54(8):774–781.

25. Vana KD, Silva GE, Goldberg R. Predictive abilities of the STOP-Bang and Epworth Sleepiness Scale in identifying sleep clinic patients at high risk for obstructive sleep apnea. *Res Nurs Health*. 2013;36(1):84–94.

26. Marti-Soler H, Hirotsu C, Marques-Vidal P, et al. The NoSAS score for screening of sleep-disordered breathing: a derivation and validation study. *Lancet Respir Med*. 2016;4(9):742–748.

27. Geay C, Mcnally S, Telhaj S. Non-native speakers of English in the classroom: what are the effects on pupil performance? *Econ J*. 2013;123(570):F281–F307.

28. Scarlata S, Pedone C, Curcio G, et al. Pre-polysomnographic assessment using the Pittsburgh Sleep Quality Index questionnaire is not useful in identifying people at higher risk for obstructive sleep apnea. *J Med Screen*. 2013;20(4):220–226.

29. Kiciński P, Przybylska-Kuć SM, Tatara K, et al. Reliability of the Epworth Sleepiness Scale and the Berlin Questionnaire for screening obstructive sleep apnea syndrome in the context of the examination of candidates for drivers. *Med Pr*. 2016;67(6):721–728.

30. Goodman E, Hinden BR, Khandelwal S. Accuracy of teen and parental reports of obesity and body mass index. *Pediatrics*. 2000;106(1 Pt 1):52–58.

31. Tuomilehto HP, Seppä JM, Partinen MM, et al; Kuopio Sleep Apnea Group. Lifestyle intervention with weight reduction: first-line treatment in mild obstructive sleep apnea. *Am J Respir Crit Care Med*. 2009;179(4):320–327.

32. Canova CR, Downs SH, Knoblauch A, Andersson M, Tamm M, Leuppi JD. Increased prevalence of perennial allergic rhinitis in patients with obstructive sleep apnea. *Respiration*. 2004;71(2):138–143.

33. Brown LK. Mild obstructive sleep apnea syndrome should be treated. Pro. *J Clin Sleep Med*. 2007;3(3):259–262.

34. Littner MR. Mild obstructive sleep apnea syndrome should not be treated. Con. *J Clin Sleep Med*. 2007;3(3):263–264.

Supplementary material
Translation for the German version of the pictorial Sleepiness and Sleep Apnoea Scale

Three German native speakers with very good knowledge of the English language independently translated the text of the English version of the pictorial Sleepiness and Sleep Apnoea Scale. The individual translations were then reviewed by DM. Differences in translation were then addressed at a consensus meeting, and an easily understandable style and wording was chosen that reflects the meaning of the original version as closely as possible. In the German version, a third answer option "don't know" represented by a question mark besides "breathing normally when asleep" and "stopping breathing when asleep" for question number 8 was added.

Scoring of sleep studies

Patients investigated at the Imperial Sleep Centre at Charing Cross Hospital, London, and Barmelweid Clinic, Barmelweid, underwent either polysomnography or polygraphy, whereas all patients investigated at the University Hospital in Basel underwent polysomnography. In all centers, Embla/Embletta Systems (Embla Systems, Broomfield, CO, USA) were used. Experienced sleep technicians or physicians scored sleep studies. Apnea was defined as an airflow limitation of >90% of baseline. Hypopnea was defined as a reduction of airflow by at least 30% and followed by saturation drop of at least 4% from baseline during at least 10 seconds.