The NoSAS score for screening of sleep-disordered breathing: a derivation and validation study

Helena Marti-Soler, Camila Hirotsu, Pedro Marques-Vidal, Peter Vollenweider, Gérard Waeber, Martin Preisig, Mehdi Tafti, Sergio Brasil Tufik, Lia Bittencourt, Sergio Tufik, José Haba-Rubio, Raphael Heinzer**

Summary

Background **Diagnosis of sleep-disordered breathing requires overnight recordings, such as polygraphy or polysomnography. Considering the cost and low availability of these procedures, preselection of patients at high risk is recommended. We aimed to develop a screening tool allowing identification of individuals at risk of sleep-disordered breathing.**

Methods **We used the participants from the population-based HypnoLaus cohort in Lausanne, Switzerland, who had a clinical assessment and polysomnography at home, to build a clinical score (the NoSAS score) using multiple factor analysis and logistic regression to identify people likely to have clinically significant sleep-disordered breathing. The NoSAS score was externally validated in an independent sleep cohort (EPISONO). We compared its performance to existing screening scores (STOP-Bang and Berlin scores).**

Findings **We used the 2121 participants from the HypnoLaus cohort who were assessed between Sept 1, 2009, and June 30, 2013. The NoSAS score, which ranges from 0 to 17, allocates 4 points for having a neck circumference of more than 40 cm, 3 points for having a body-mass index of 25 kg/m² to less than 30 kg/m² or 5 points for having a body-mass index of 30 kg/m² or more, 2 points for snoring, 4 points for being older than 55 years of age, and 2 points for being male. Using a threshold of 8 points or more, the NoSAS score identified individuals at risk of clinically significant sleep-disordered breathing, with an area under the curve (AUC) of 0·74 (95% CI 0·72–0·76). It showed an even higher performance in the EPISONO cohort, with an AUC of 0·81 (0·77–0·85). The NoSAS score performed significantly better than did the STOP-Bang (AUC 0·67 [95% CI 0·65–0·69]; p<0·0001) and Berlin (0·63 [0·61–0·66]; p<0·0001) scores.**

Interpretation **The NoSAS score is a simple, efficient, and easy to implement score enabling identification of individuals at risk of sleep-disordered breathing. Because of its high discrimination power, the NoSAS score can help clinicians to decide which patients to further investigate with a nocturnal recording.**

Funding **Faculty of Biology and Medicine of the University of Lausanne, Lausanne University Hospital, Swiss National Science Foundation, Leenaards Foundation, GlaxoSmithKline, and Vaud Pulmonary League.**

Introduction

Sleep-disordered breathing is a highly prevalent disease, characterised by repetitive collapse of the upper airway during sleep. Moderate-to-severe sleep-disordered breathing has been shown to affect half of men and almost a quarter of women in the general middle-to-old-age population (40–85 years old).[1] The disorder affects quality of life and different aspects of health domains;[2,3] is associated with an increased incidence of cardiovascular events,[4,5] hypertension,[6] diabetes,[7] metabolic syndrome,[8] and car accidents;[9] and represents an important public health issue. Diagnosis requires overnight recordings with portable limited-channel recorders (respiratory polygraphy) or full polysomnography (PSG), either at home or in a sleep laboratory. These investigations are time-consuming and expensive and cannot be used as routine screening.

Different clinical scores, such as the Berlin[10] or STOP-Bang[11] scores, have therefore been previously proposed as screening tools. The Berlin Questionnaire[10] includes information about snoring, daytime sleepiness or fatigue, obesity, and hypertension. It was created with use of a general clinical practice sample of 744 individuals, of whom 100 (13%) had a polygraphic recording at home to confirm the diagnosis. The STOP-Bang score[11] combines information from a self-administered questionnaire about complaints of snoring, tiredness, observed apnoea, and high blood pressure with clinical characteristics such as body-mass index (BMI), age, neck circumference, and sex. It was created based on a large group of patients assessed before a surgical procedure (2477 patients), of whom 211 (9%) had a full PSG to confirm the diagnosis.

These instruments are used in current clinical practice, although they were developed with use of less sensitive recording technology (the Berlin score was developed with thermistors) or older respiratory event scoring criteria (STOP-Bang) than those used nowadays, which therefore do not match current standards. Considering that these technical differences have been shown to have a substantial impact on sleep-disordered breathing diagnosis and perceived prevalence,[1] we aimed to develop a new screening tool for sleep-disordered breathing using current recording standards as a reference and external validation to ensure its reliability in different populations.

Lancet Respir Med 2016

Published Online
June 16, 2016
http://dx.doi.org/10.1016/S2213-2600(16)30075-3

See Online/Comment
http://dx.doi.org/10.1016/S2213-2600(16)30119-9

*Contributed equally

Institute of Social and Preventive Medicine (H Marti-Soler PhD) **and Center for Integrative Genomics** (Prof M Tafti PhD), **University of Lausanne, Lausanne, Switzerland; Departamento de Psicobiologia, Universidade Federal de São Paulo, São Paulo, Brazil** (C Hirotsu PhD, L Bittencourt PhD, Prof S Tufik PhD); **Department of Internal Medicine** (P Marques-Vidal MD, Prof P Vollenweider MD, Prof G Waeber MD), **Psychiatry Department** (Prof M Preisig MD), **Center for Investigation and Research in Sleep** (Prof M Tafti, J Haba-Rubio MD, R Heinzer MD), **and Pulmonary Department** (R Heinzer), **University Hospital of Lausanne, University of Lausanne, Lausanne, Switzerland; and Medical School, Universidade de São Paulo, São Paulo, Brazil** (S B Tufik MD)

Correspondence to:
Dr José Haba-Rubio or Dr Raphael Heinzer, Center for Investigation and Research in Sleep, University Hospital of Lausanne, University of Lausanne, 1011 Lausanne, Switzerland
jose.haba-rubio@chuv.ch or **raphael.heinzer@chuv.ch**

Research in context

Evidence before this study

Sleep-disordered breathing is a highly prevalent condition associated with neurocognitive deficits and increased cardiovascular risk. Diagnostic procedures, such as respiratory polygraphy or polysomnography, are time-consuming and expensive, so they cannot be used as routine screening for sleep-disordered breathing. We searched PubMed with the terms "sleep apnea", "sleep apnoea", and "screening score" up to Nov 30, 2015 (in English). We found that the most commonly used scores for sleep apnoea screening in clinical settings were the Berlin and the STOP-Bang scores. However, clinicians often complain about the complexity of the Berlin score and the high positive rate and lack of discrimination of the STOP BANG score. We thus aimed to develop a new screening score for sleep apnoea based on HypnoLaus, a large population-based sleep cohort.

Added value of this study

We propose a new simple clinical score (the NoSAS score) to screen for sleep-disordered breathing on the basis of clinical and polysomnographic data from a large population-based cohort. Using variables easily available in primary care practice, the NoSAS score allows identification of individuals at risk and ruling out of clinically significant sleep-disordered breathing, with an NPV of 90% and 98% in two ethnically different population-based cohorts.

Implications of all the available evidence

Compared with existing screening scores, the NoSAS score helps clinicians decide with a higher accuracy than at present which individuals should be referred for further testing, thereby reducing the number of missed sleep-disordered breathing diagnoses and unnecessary nocturnal recordings.

Methods

Study design and participants

The HypnoLaus Sleep Cohort study[1] included a random subset of the population-based CoLaus/PsyCoLaus cohort[12,13] who had a full PSG at home in Lausanne, Switzerland. Sleep-related complaints and habits were investigated with use of the Pittsburgh Sleep Quality Index[14] and Epworth Sleepiness Scale,[15] and the probability of sleep-disordered breathing was investigated using the Berlin Questionnaire[10] and STOP-Bang[11] score. Clinical and demographic data were also collected (appendix). The ethics committee of the University of Lausanne approved the CoLaus/PsyCoLaus cohort study and the HypnoLaus Sleep Cohort study. Written informed consent was obtained from all participants.

See Online for appendix

Procedures

In the HypnoLaus Sleep Cohort study, certified technicians equipped participants with a PSG recorder (Titanium; Embla, Flaga, Reykjavik, Iceland) between 1700 h and 2000 h at the Center for Investigation and Research in Sleep, University Hospital of Lausanne (Lausanne, Switzerland). All sleep recordings took place in patients' home environments and in accordance with the American Academy of Sleep Medicine's 2007 recommended setup specifications.[16] Breathing was recorded with nasal pressure sensors. Participants being treated for sleep-disordered breathing were asked to discontinue their treatment 1 week before the sleep recording. Two trained sleep technicians masked to questionnaire results manually scored PSG recordings using Somnologica (Embla) software (version 5.1.1). Each recording was reviewed by an expert sleep physician, and a second sleep expert did random quality checks. Quality control for concordance between the two PSG scorers was implemented periodically to ensure that both scorers achieved at least a 90% level of agreement for sleep stages and respiratory events and an 85% level of agreement for arousals.[17] Apnoea was defined as a decrease of at least 90% of airflow from baseline, lasting 10 s or longer. Hypopnoeas were scored with use of the latest 2012 American Academy of Sleep Medicine criteria[18] (≥30% decrease of airflow lasting at least 10 s, associated with either an arousal or a ≥3% O_2 saturation decrease). The mean number of apnoeas and hypopnoeas per h of sleep (apnoea-hypopnoea index [AHI]) was calculated. We defined clinically significant sleep-disordered breathing as an AHI of more than 20 events per h according to the initial analysis of the HypnoLaus Sleep Cohort.[1]

We used the EPISONO population-based cohort,[19] studied in São Paulo, Brazil, in 2007, to externally validate the NoSAS score. The details of the study design and methods have been previously published.[20] Briefly, the EPISONO survey used a probabilistic three-stage cluster sample of São Paulo inhabitants to represent the population according to sex, age, and socioeconomic status. Face-to-face interviews and in-lab full-night PSG with a nasal cannula and thermistor were done. An initial sample of 1101 participants was included from the four homogeneous socioeconomic macroregions of São Paulo, as classified by the National Institute of Demographics and Statistics census. 59 (5%) refused to undergo the PSG examination, leading to a final sample size of 1042 (95%) participants. Apnoeas were defined as a decrease of at least 90% of airflow from baseline, lasting 10 s or longer. Hypopnoeas were defined as a 50% or higher decrease of airflow lasting at least 10 s, associated with either an arousal or a 3% or higher O_2 saturation decrease.

Statistical analysis

To develop the screening score, we used multiple factor analysis based on demographic data and clinical charac-

teristics to visually explore the simultaneous relationships between continuous and categorical variables and identify sleep-disordered breathing-related features. For each selected continuous variable, we selected the threshold with the highest area under the curve (AUC) using an AHI of more than 20 or less than or equal to 20 as the response variable. We estimated the probability of clinically significant sleep-disordered breathing (an AHI of >20 events per h) using logistic regression for participants with complete data and created a score taking into account the contribution of each variable included in the logistic model. We built the score as the sum of assigned points, defined as the exponential of the estimated coefficient, rounded to the nearest integer. We subsequently dichotomised it into positive or negative.

We assessed the performance of the score in terms of discrimination using the AUC and DeLong method[21] to calculate 95% CIs, positive predictive value (PPV; the probability that participants with a positive screening test truly have the disease), and negative predictive value (NPV; the probability that participants with a negative screening test truly do not have the disease). To avoid overfitting due to use of the same data for estimation and evaluation, we used the k-fold cross validation technique,[22] with k equal to 10. Briefly, this technique randomly splits data into k equal size subsamples. It uses k–1 subsamples for estimation and the remaining one for validation. The process is then repeated k times, with each of the k subsamples used once as the validation data. The k results are then combined to produce a single estimation. We calculated the percentage of unnecessary sleep recordings (false positives) and missed sleep-disordered breathing diagnoses (false negatives) by comparing the results of our score with whether their polysomnography showed an AHI of more or less than or equal to 20 events per h. We did statistical analyses using R software (version 3.1.0). We did an independent external validation using data from the EPISONO population-based cohort.[19] The performance of our score, as well as the Berlin and STOP-Bang scores, were assessed and compared with one another in the EPISONO sample.

Role of the funding source

The Faculty of Biology and Medicine of the University of Lausanne, Lausanne University Hospital, the Leenaards Foundation, and the Vaud Pulmonary League funded the salary of the technicians who did the sleep recordings. The Swiss National Science Foundation funded the statisticians and supported the initial CoLaus cohort. GlaxoSmithKline supported the initial CoLaus cohort and funded the PSG recorders. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding authors had full access to all of the data in the study and had final responsibility for the decision to submit for publication.

Results

Of the 3043 participants contacted from the CoLaus/PsyCoLaus study, 2168 (71%) agreed to have a PSG at home in the HypnoLaus study, assessed between Sept 1, 2009, and June 30, 2013. Among the 2168 participants who had a full PSG at home, 2121 (98%) were included in the HypnoLaus study and 47 (2%) were excluded (41 because of a total sleep time of <4 h and six because of technical problems). Compared with the whole CoLaus/PsyCoLaus cohort, individuals who had PSG in the HypnoLaus study were similar in terms of age, sex, BMI, and ethnic origin,

	HypnoLaus cohort (n=2121)
Age (years)	59 (11)
Male	1024 (48%)
Overweight (body-mass index >25 kg/m²)	1210 (57%)
Body-mass index (kg/m²)	25·6 (4·1)
Neck circumference (cm)*	36·9 (3·9)
Waist-to-hip ratio†	0·9 (<0·1)
Use alcohol‡	560 (26%)
Current smoker	395 (19%)
Snore§	1164/1759 (66%)
Hypertension¶	877 (41%)
Diabetes‖	212 (10%)
Metabolic syndrome**	641 (30%)
Apnoea-hypopnoea index >20/h	551 (26%)
Epworth score	6 (3–9)
Epworth score >10	258 (12%)
PSQI score	4 (3–7)
Berlin score ≥2	525 (25%)
STOP-Bang score ≥3	1076/1559 (69%)

Data are mean (SD), n (%), or median (IQR). PSQI=Pittsburgh Sleep Quality Index. *Measured between mid-cervical spine and mid-anterior neck, just below the laryngeal prominence. †Calculated as the ratio of the circumference of the waist to that of the hip. ‡Self-reported use of alcohol (at least one drink) in the evening preceding the polysomnography. §Some data missing. ¶Defined as a systolic blood pressure of 140 mm Hg or higher or a diastolic blood pressure of 90 mm Hg or higher, or current use of antihypertensive medication. ‖Defined as a fasting blood glucose concentration of 7 mmol/L or higher or current use of antidiabetic medication. **Defined according to the Adult Treatment Panel III report.[23]

Table 1: **Demographic data and characteristics of the HypnoLaus cohort**

	Points
Neck circumference >40 cm	4
Obesity	
BMI 25 kg/m² to <30 kg/m²	3
BMI ≥30 kg/m²	5
Snoring	2
Age >55 years	4
Sex: male	2

The patient has a high probability of sleep-disordered breathing if they have a NoSAS score of 8 or higher. BMI=body-mass index.

Table 2: **NoSAS score**

and were considered representative of Lausanne's general population.[12] Demographic data and characteristics of the 2121 HypnoLaus participants are shown in table 1.

38 (2%) participants were being treated for sleep-disordered breathing. Initially, we analysed all available demographic and clinical variables classically related to sleep-disordered breathing, such as age, sex, BMI, height, weight, neck circumference, waist-to-hip ratio, alcohol use, smoking, snoring, hypertension, diabetes, metabolic syndrome, and Epworth score, with multiple factor analysis to identify blocks of variables. Using a graphical tool provided by this approach and logistic regression models, we have identified the most relevant variables: neck circumference, obesity, snoring, age, and sex (NoSAS score). This score, which ranges from 0 to 17, allocates 4 points for having a neck circumference of more than 40 cm, 3 points for having a BMI of 25 kg/m² to less than 30 kg/m² or 5 points for having a BMI of 30 kg/m² or more, 2 points for snoring, 4 points for being older than 55 years of age, and 2 points for being male (table 2).

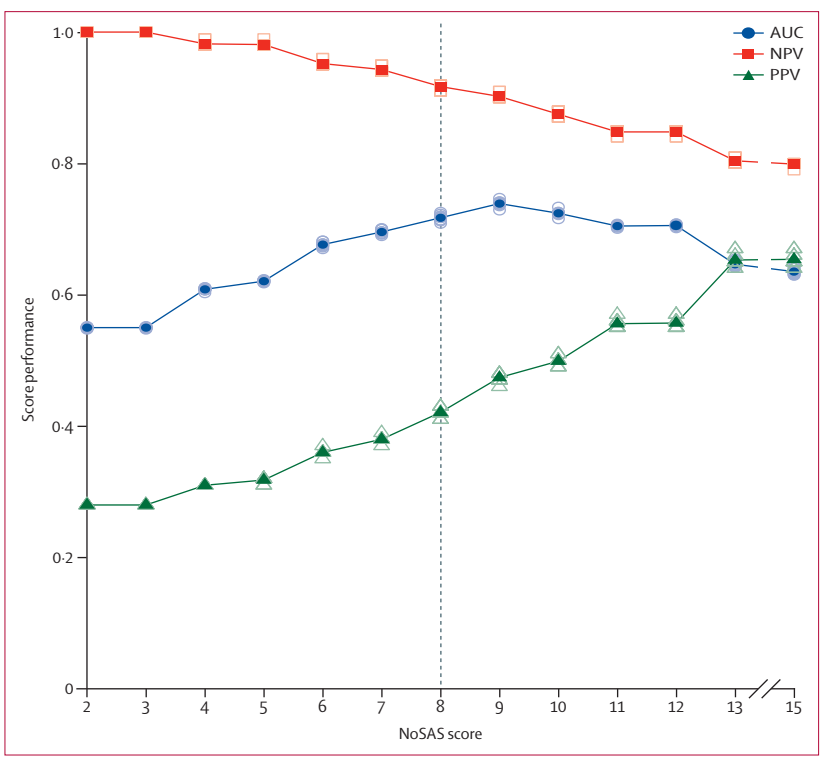

Figure 1: **Performance of the NoSAS score for each possible threshold**
The multiple data points represent the ten-fold cross-validation method. AUC=area under the curve. NPV=negative predictive value. PPV=positive predictive value. The dotted line shows the threshold of 8. Values of 1, 14, and 16 do not appear as no NoSAS scores are possible for these values, and values of 0 and 17 do not appear because they represent 0 and 1 by definition.

We chose 8 as a threshold, so the score was defined as positive if it was greater than or equal to 8 points and negative if it was fewer than 8 points, on the basis of its ability to discriminate participants with clinically significant sleep-disordered breathing (ie, an AHI of >20 events per h), its PPV, and its NPV (figure 1). The score had an AUC of 0·74 (95% CI 0·72–0·76), a PPV of 0·47 (0·44–0·51), and an NPV of 0·90 (0·88–0·92; table 3). We also compared the performance of the NoSAS score with that of the STOP-Bang and Berlin scores using AHI cutoffs of five, 10, 15, 20, 25, and 30 events per h (figure 2, appendix). Overall, the NoSAS score was able to detect 1666 (98%) of 1703 participants with severe sleep-disordered breathing (an AHI of >30 events per h), with 37 (2%) of 1703 having a false-negative diagnosis. STOP-Bang correctly classified 839 (54%) of 1559 participants for a cutoff of 20 events per h and 683 (44%) of 1559 participants for severe sleep-disordered breathing.

In addition to the internal cross-validation, the NoSAS score was also validated in the independent population-based EPISONO cohort[19] (466 [45%] of 1042 were men; mean age 42 years (SD 14); 621 [60%] of 1042 were overweight). The NoSAS score had an AUC of 0·81 (95% CI 0·77–0·85), a PPV of 0·33 (0·28–0·38), and an NPV of 0·98 (0·96–0·99). Comparing the classification capability of the NoSAS, STOP-Bang, and Berlin scores, the NoSAS score performed significantly better than did the other scores in both cohorts. In HypnoLaus, the AUC was 0·74 (0·72–0·76) for the NoSAS score, 0·67 (0·65–0·69; p<0·0001) for the STOP-Bang score, and

	Estimated prevalence of sleep-disordered breathing (n/N [%])	AUC	Sensitivity	Specificity	PPV	NPV
HypnoLaus						
NoSAS	737/1703 (43%)	0·74 (0·72–0·76)	0·79 (0·75–0·82)	0·69 (0·67–0·72)	0·47 (0·44–0·51)	0·90 (0·88–0·92)
STOP-Bang	1076/1559 (69%)	0·67 (0·65–0·69)	0·94 (0·91–0·96)	0·40 (0·37–0·43)	0·35 (0·33–0·39)	0·95 (0·93–0·97)
Berlin	525/2115 (25%)	0·63 (0·61–0·66)	0·41 (0·37–0·45)	0·81 (0·79–0·83)	0·43 (0·39–0·47)	0·79 (0·77–0·81)
EPISONO						
NoSAS	310/1042 (30%)	0·81 (0·77–0·85)	0·85 (0·77–0·91)	0·77 (0·75–0·80)	0·33 (0·28–0·38)	0·98 (0·96–0·99)
STOP-Bang	411/1016 (40%)	0·68 (0·63–0·73)	0·72 (0·63–0·80)	0·64 (0·61–0·67)	0·21 (0·17–0·25)	0·95 (0·92–0·96)
Berlin	362/1036 (35%)	0·65 (0·59–0·70)	0·61 (0·52–0·70)	0·68 (0·65–0·71)	0·20 (0·16–0·24)	0·93 (0·91–0·95)

Data in parentheses are 95% CIs. AUC=area under the curve. PPV=positive predictive value. NPV=negative predictive value.

Table 3: **Performance of different scores in the HypnoLaus and EPISONO cohorts**

0·63 (0·61–0·66; p<0·0001) for the Berlin score. In EPISONO, the AUC was 0·81 (0·77–0·85) for the NoSAS score, 0·68 (0·63–0·73; p<0·0001) for the STOP-Bang score, and 0·65 (0·59–0·70; p<0·0001) for the Berlin score (table 3). The proportion of correct sleep-disordered breathing classifications—ie, positive NoSAS score when the participant has an AHI of more than 20 events per h and negative NoSAS score when they have an AHI of 20 events per h or fewer, in the HypnoLaus and EPISONO cohorts, is shown in table 4. All participants who had a false-negative result with the NoSAS score had a neck circumference of 40 cm or lower, 33 (35%) of 95 had a BMI between 25 kg/m² and 30 kg/m², and only four (4%) of 95 had a BMI of 30 kg/m² or higher.

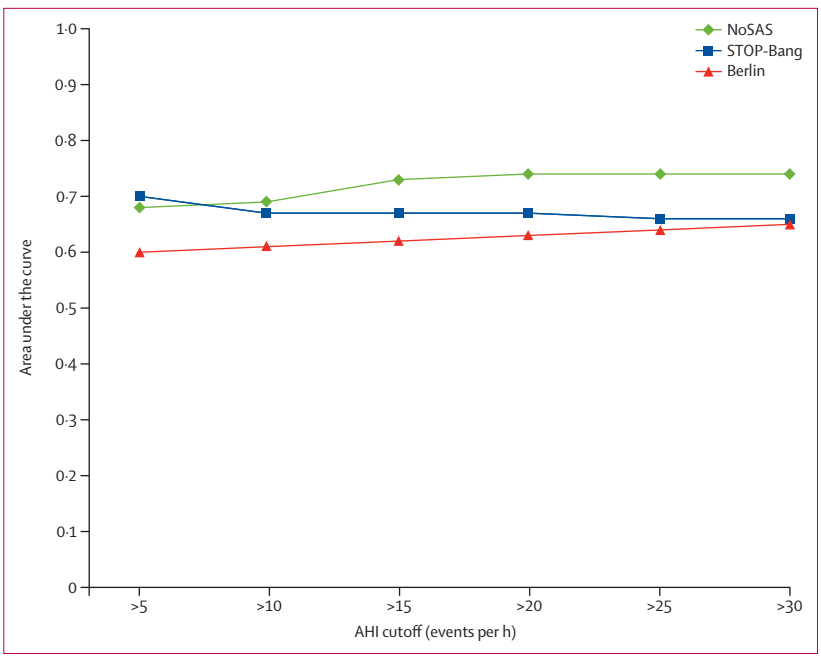

Figure 2: **Performance of the NoSAS score compared with STOP-Bang and Berlin scores**
AHI=apnoea-hypopnoea index.

Discussion

Using a large population-based cohort, we propose a new simple and easy-to-interpret clinical score to screen for clinically significant sleep-disordered breathing. Our results show that—using variables easily available in primary care practice—the NoSAS score allows clinically significant sleep-disordered breathing to be reliably ruled out, with an NPV of 90% and 98% in two ethnically different population-based cohorts. Compared with existing screening scores, the NoSAS score allows a reduction in the number of unnecessary nocturnal recordings as well as the number of missed diagnoses of sleep-disordered breathing.

An ideal sleep-disordered breathing screening score should have a high sensitivity to avoid false-negative results, but also be specific enough to avoid referral of low-risk patients for costly and time-consuming sleep recordings. Although the proportion of correct classifications of the Berlin Questionnaire was only mildly lower than that of the NoSAS score, its sensitivity was quite low in both cohorts, yielding the highest proportion of missed diagnoses, which is unfortunate for a screening score. As suggested by the higher AUC and correct classification proportion than the other scores, the NoSAS score seems to represent the best sensitivity specificity compromise, allowing clinically significant sleep-disordered breathing to be reliably ruled out without yielding too many unnecessary sleep investigations.

Overall, the NoSAS score had a false-negative proportion of 2% for participants with severe sleep-disordered breathing. Although the proportion of missed diagnoses with the NoSAS score with use of different thresholds is much lower than with the Berlin score, it is higher than with the STOP-Bang score in the HypnoLaus cohort (but not in the EPISONO cohort). This very low false-negative proportion of the STOP-Bang score in the HypnoLaus cohort is, however, obtained at the cost of a very high false-positive proportion (45%). Moreover, the ability of STOP-Bang to correctly classify participants for a cutoff of 20 events per h is close to a random classification and worse than random for severe sleep-disordered breathing, which questions its use in these sleep-disordered breathing categories. Participants with a false-negative result with the NoSAS score tended to have a low neck circumference and BMI. The cause of their sleep-disordered breathing is therefore more likely to be related to maxillofacial deformities, a high loop gain, or upper airway muscle control dysfunction than to obesity.

The fact that the NoSAS score performed consistently well in two unrelated samples with different age ranges, ethnic backgrounds, and habits is an asset and strongly emphasises its consistency and generalisability to different populations. Conversely, STOP-Bang seems to be less consistent[24] since it showed a very low specificity to predict moderate-to-severe and severe sleep-disordered breathing in some population samples.[25,26] We also observed variations in STOP-Bang and Berlin performances between the HypnoLaus and EPISONO cohorts, whereas the performance of the NoSAS score appeared to be more stable than were the other scores. This higher stability could be due to the lower number of subjective variables included in the NoSAS score.

To be time efficient, a screening score should include a small number of items related to easily available and objective variables. The NoSAS score uses biometric items, such as age, sex, and BMI, which are part of any standard clinical assessment, and neck circumference, which can be easily measured. As the aim of this analysis was to develop a reliable and easily applicable score, we tried to limit the number of subjective variables, such as witnessed sleep apnoeas, which require the observation of a bed partner, and subjective sleepiness, which was not associated with sleep-disordered breathing severity in our cohort.[1] This absence of association could be due

	NoSAS		STOP-Bang		Berlin	
	Positive	Negative	Positive	Negative	Positive	Negative
HypnoLaus (NoSAS n=1703; STOP-Bang n=1559; Berlin n=2115)*						
AHI ≤20 events per h (negative)	388 (23%)†	871 (51%)‡	697 (45%)†	460 (30%)‡	300 (14%)†	1262 (60%)‡
AHI >20 events per h (positive)	349 (20%)‡	95 (6%)§	379 (24%)‡	23 (1%)§	225 (11%) ‡	328 (16%)§
EPISONO (NoSAS n=1042; STOP-Bang n=1016; Berlin n=1036)¶						
AHI ≤20 events per h (negative)	208 (20%)†	714 (69%)‡	326 (32%)†	572 (56%)‡	290 (28%)†	628 (61%)‡
AHI >20 events per h (positive)	102 (10%)‡	18 (2%)§	85 (8%)‡	33 (3%)§	72 (7%)‡	46 (4%)§

Data are n (%). AHI=apnoea-hypopnoea index. *Prevalence of sleep-disordered breathing (AHI >20 events per h) by polysomnography was 26%. †Unnecessary sleep recording (false positive). ‡Correct classification. §Missed diagnosis (false negative). ¶Prevalence of sleep-disordered breathing (AHI >20 events per h) by polysomnography was 12%.

Table 4: **Correct classification, unnecessary sleep recordings, and missed diagnosis according to the HypnoLaus and EPISONO cohorts**

to the multiple factors that influence reported sleepiness, such as sleep duration, medications, mood disorders, or other sleep disorders. We did not include hypertension as it did not appear to be a major independent predictor of sleep-disordered breathing in our cluster analysis. This absence could also be due to the multiple factors that can influence blood pressure. The presence of snoring is the only subjective item we included in the NoSAS score because of its strong statistical association with clinically significant sleep-disordered breathing. However, we did not include snoring severity and frequency since these variables are more subjective and difficult to quantify than is presence of snoring.

Compared with the eight items of STOP-Bang and the nine to eleven questions (depending on the snore question) and three separate calculations required for the Berlin Questionnaire, we believe that the NoSAS score, which includes only five items, is easier to use in clinical practice than are the other two scores. Additionally, a free smartphone application is available (NoSAS Score; BeFine, Geneva, Switzerland). An online NoSAS calculator will also be available. We can therefore hope that this easy-to-use score could encourage general practitioners to screen for sleep-disordered breathing and decrease the proportion of undiagnosed and untreated people with sleep-disordered breathing.

One of the strengths of our analysis is that it relies on a large population-based sample, with more than 2000 polysomnographic recordings, whereas a much lower number of polysomnographies than used in this study was used as a reference to develop the Berlin[10] and STOP-Bang[11] scores. Moreover, we used up-to-date recording techniques and scoring criteria currently used in sleep laboratories, which probably increases the relevance of the NoSAS score for current clinical practice. Furthermore, the external validation in an unrelated cohort from another continent showing high reproducibility of the NoSAS score is an asset. Limitations to our analysis also exist that need to be taken into account. First, we chose an AHI of more than 20 events per h as a cutoff to build the score on the basis of the results of the first analysis of the HypnoLaus cohort showing significant associations between sleep-disordered breathing and hypertension, diabetes, metabolic syndrome, and depression, almost exclusively in the upper quartile (an AHI of >20·6 events per h).[1] This decision was clearly arbitrary, but the NoSAS score seems to also perform well with other AHI cutoffs. When thresholds of five, 10, 15, 20, 25, and 30 events per h were used, we saw that the performance of the NoSAS score remained higher than with the other scores for AHI thresholds of ten events per h or more. Since the currently recommended recording techniques and scoring criteria for respiratory events are highly sensitive, yielding high sleep-disordered breathing prevalence figures,[27] we believe that use of an AHI threshold of less than 10 events per h would not be clinically relevant. Second, some data were missing for the calculation of each of the scores because some participants did not fill in the questionnaires properly. This proportion was, however, rather low in general, and we do not believe that it had a substantial impact on the results because we are not in the framework of a point estimate like a treatment or an exposure effect, for instance. Finally, the NoSAS score was built and validated in population-based cohorts, which might differ from clinical populations. This score will therefore have to be further validated in specific clinical samples.

Contributors

HM-S, JH-R, and RH wrote the manuscript and interpreted results. HM-S, CH, and SBT analysed data. All authors acquired data and critically revised the manuscript for important intellectual content. JH-R and RH supervised the study.

Declaration of interests

We declare no competing interests.

Acknowledgments

The HypnoLaus, CoLaus, and PsyCoLaus studies were supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of the University of Lausanne, the Swiss National Science Foundation (grants 3200B0–105993, 3200B0–118308, 33CSCO-122661, 33CS30–139468, and 33CS30–148401), the Leenaards Foundation, and the Vaud Pulmonary League (Ligue Pulmonaire Vaudoise). We thank the Lausanne population who volunteered to participate in the CoLaus, PsyCoLaus, and HypnoLaus studies and the whole team of CoLaus.

References

1 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:** 310–18.

2 Jordan AS, McSharry DG, Malhotra A. Adult obstructive sleep apnoea. *Lancet* 2014; **383:** 736–47.

3 Malhotra A, White DP. Obstructive sleep apnoea. *Lancet* 2002; **360:** 237–45.

4 Marin JM, Carrizo SJ, Vicente E, Agusti AG. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet* 2005; **365:** 1046–53.

5 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:** 2034–41.

6 Peppard PE, Young T, Palta M, Skatrud J. Prospective study of the association between sleep-disordered breathing and hypertension. *N Engl J Med* 2000; **342:** 1378–84.

7 Kendzerska T, Gershon AS, Hawker G, Tomlinson G, Leung RS. Obstructive sleep apnea and incident diabetes. A historical cohort study. *Am J Respir Crit Care Med* 2014; **190:** 218–25.

8 Coughlin SR, Mawdsley L, Mugarza JA, Calverley PM, Wilding JP. Obstructive sleep apnoea is independently associated with an increased prevalence of metabolic syndrome. *Eur Heart J* 2004; **25:** 735–41.

9 Teran-Santos J, Jimenez-Gomez A, Cordero-Guevara J. The association between sleep apnea and the risk of traffic accidents. Cooperative Group Burgos-Santander. *N Engl J Med* 1999; **340:** 847–51.

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

11 Chung F, Yegneswaran B, Liao P, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008; **108:** 812–21.

12 Firmann M, Mayor V, Vidal PM, et al. The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovasc Disord* 2008; **8:** 6.

13 Preisig M, Waeber G, Vollenweider P, et al. The PsyCoLaus study: methodology and characteristics of the sample of a population-based survey on psychiatric disorders and their association with genetic and cardiovascular risk factors. *BMC Psychiatry* 2009; **9:** 9.

14 Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989; **28:** 193–213.

15 Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991; **14:** 540–45.

16 Iber C, Ancoli-Israel S, Chesson AL, Qaun SF. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. Westchester: American Academy of Sleep Medicine, 2007.

17 Redline S, Sanders MH, Lind BK, et al. Methods for obtaining and analyzing unattended polysomnography data for a multicenter study. Sleep Heart Health Research Group. *Sleep* 1998; **21:** 759–67.

18 Berry RB, Budhiraja R, Gottlieb DJ, et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 2012; **8:** 597–619.

19 Tufik S, Santos-Silva R, Taddei JA, Bittencourt LR. Obstructive sleep apnea syndrome in the Sao Paulo Epidemiologic Sleep Study. *Sleep Med* 2010; **11:** 441–46.

20 Santos-Silva R, Tufik S, Conway SG, Taddei JA, Bittencourt LR. Sao Paulo Epidemiologic Sleep Study: rationale, design, sampling, and procedures. *Sleep Med* 2009; **10:** 679–85.

21 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; **44:** 837–45.

22 Stone M. Cross-validatory choice and assessment of statistical predictions. *J Royal Stat Soc Series B* 1974; **36:** 111–47.

23 National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation* 2002; **106:** 3143–421.

24 Miller JN, Berger AM. Screening and assessment for obstructive sleep apnea in primary care. *Sleep Med Rev* 2015; **29:** 41–51.

25 Silva GE, Vana KD, Goodwin JL, Sherrill DL, Quan SF. Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. *J Clin Sleep Med* 2011; **7:** 467–72.

26 El-Sayed I. Comparison of four sleep questionnaires for screening obstructive sleep apnoea. *Egypt J Chest Dis Tuberc* 2012; **61:** 433–41.

27 Heinzer R, Marti-Soler H, Haba-Rubio J. Prevalence of sleep apnoea syndrome in the middle to old age general population. *Lancet Respir Med* 2016; **4:** e5–6.