pii: jc-17-00556 http://dx.doi.org/10.5664/jcsm.7202

JCSM
Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

Simplifying the Screening of Obstructive Sleep Apnea With a 2-Item Model, No-Apnea: A Cross-Sectional Study

Ricardo L.M. Duarte, MD, MSc[1,2]; Marcelo F. Rabahi, MD, PhD[3]; Flavio J. Magalhães-da-Silveira, MD[1]; Tiago S. de Oliveira-e-Sá, MD[4,5]; Fernanda C.Q. Mello, MD, PhD[2]; David Gozal, MD, MBA[6]

[1]Sleep - Laboratório de Estudo dos Distúrbios do Sono, Centro Médico BarraShopping, Rio de Janeiro, Brazil; [2]Instituto de Doenças do Tórax - Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; [3]Faculdade de Medicina, Universidade Federal de Goiás, Goiás, Brazil; [4]Hospital de Santa Marta - Centro Hospitalar Lisboa Central, Portugal; [5]NOVA Medical School - Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Portugal; [6]Department of Pediatrics, Pritzker School of Medicine, Biological Sciences Division, The University of Chicago, Chicago, Illinois

Study Objectives: To develop and validate a practical model for obstructive sleep apnea (OSA) screening in adults based on objectively assessed criteria, and then compare it with two widely used tools, namely STOP-BANG and NoSAS.

Methods: This is a retrospective study of an existing database of consecutive outpatients who were referred for polysomnography for suspected sleep-disordered breathing by their primary care physicians. Area under the curve (AUC) and 2 × 2 contingency tables were employed to obtain the performance of the new instrument.

Results: A total of 4,072 subjects were randomly allocated into two independent cohorts: one for derivation (n = 2,037) and one for validation (n = 2,035). A mnemonic model, named No-Apnea, with two variables (neck circumference and age) was developed (total score: 0–9 points). We used the cutoff ≥ 3 to classify patients at high risk of having OSA. OSA severity was categorized by apnea-hypopnea index (AHI): any OSA (AHI 5 ≥ events/h; OSA-5), moderate/severe OSA (AHI 15 ≥ events/h; OSA-15); and severe OSA (AHI 30 ≥ events/h; OSA-30). In the derivation cohort, the AUCs for screening of OSA-5, OSA-15, and OSA-30 were: 0.784, 0.758, and 0.754; respectively. The rate of subjects correctly screened was 78.1%, 68.8%, and 54.4%, respectively for OSA-5, OSA-15, and OSA-30. Subsequently, the model was validated confirming its reproducibility. In both cohorts, No-Apnea discrimination was similar to STOP-BANG or NoSAS.

Conclusions: The No-Apnea, a 2-item model, appears to be a useful and practical tool for OSA screening, mainly when limited resources constrain referral evaluation. Despite its simplicity when compared to previously validated tools (STOP-BANG and NoSAS), the instrument exhibits similar performance characteristics.

Keywords: clinical assessment, obstructive sleep apnea, polysomnography, scoring

Citation: Duarte RL, Rabahi MF, Magalhães-da-Silveira FJ, de Oliveira-e-Sá TS, Mello FC, Gozal D. Simplifying the screening of obstructive sleep apnea with a 2-item model, No-Apnea: a cross-sectional study. *J Clin Sleep Med.* 2018;14(7):1097–1107.

BRIEF SUMMARY

Current Knowledge/Study Rationale: There are several clinical questionnaires for screening of obstructive sleep apnea (OSA), aiming to find the patients at high risk for this disease, which is often undiagnosed. These questionnaires often use subjective data related to sleep, which often requires bed partner information, reducing their practical applicability.

Study Impact: The No-Apnea, a practical model with only two variables (neck circumference and age) may be useful to identify patients at high risk for OSA, indicating home sleep studies and reducing long waiting lines found in several sleep laboratories, especially in primary care settings and when limited resources constrain referral evaluation. When compared to STOP-BANG (an 8-item model) and NoSAS (a 5-item model), the No-Apnea (a 2-item model) had similar performance characteristics.

INTRODUCTION

Obstructive sleep apnea (OSA) is characterized by frequent partial or complete collapse of the upper airway during sleep, resulting in periodic hypoxemia, increased respiratory effort, and arousals.[1] There is growing evidence of OSA playing a role in the pathogenesis of cardiovascular and metabolic diseases[1] and also a strong association with increased mortality rates.[2] Recent studies[3–5] indicate a clear increase in the prevalence of OSA: a Brazilian study showed that OSA was diagnosed in 32.8% of the participants[3]; two other studies have also shown similar findings, with a prevalence of moderate to severe OSA of 13% being recorded among men and of 6% among women[4] and of 49.7% in men and 23.4% in women.[5] Aging populations[6] and the strong association between body mass index (BMI) and OSA[7] may account for such increases in OSA prevalence. Indeed, obesity has been linked with increases in neck circumference (NC), which can reduce the upper airway diameter and alter the mechanical properties of the upper airway altogether.[8,9]

The gold standard for OSA diagnosis consists of full polysomnography (PSG); however, it is not readily available for the large number of patients with suspected OSA, such that

a screening tool could be useful to stratify patients at risk for OSA, and enable improved access to PSG testing and diagnosis.[10] Several OSA screening questionnaires have been described,[11–16] among which the Berlin and STOP-BANG instruments are most frequently cited. The Berlin questionnaire[12] includes items that access the presence and frequency of snoring, sleepiness or fatigue, and history of obesity or hypertension. The STOP questionnaire (snoring, tiredness, observed apnea, and hypertension), and the STOP-BANG questionnaire (STOP plus BMI, age, NC, and gender) were first developed and validated in surgical patients.[13] Both questionnaires are self-administered consisting of 4 or 8 yes-or-no questions, respectively. For OSA diagnosis in a sample of surgical patients, the STOP-BANG questionnaire showed the following parameters: sensitivity 83.6%, specificity 56.4%, positive predictive value (PPV) 81.0%, and negative predictive value (NPV) 60.8%.[13]

Despite the extensive efforts to develop highly performing questionnaires, objectively based scoring tools for screening of OSA have rarely been evaluated. Indeed, the existing OSA screening models rely on subjective information, such as snoring, observed apnea, and choking/gasping, usually provided by the bed partner, and therefore not always available. To overcome such constraints in the screening of OSA, our study comprised two parts: (1) development and validation of a very practical screening model and (2) comparison with STOP-BANG questionnaire[13] and the recently reported NoSAS score.[16]

METHODS

Study Design and Patient Selection

This is a cross-sectional retrospective study of an existing database from January 2015 to December 2016 of consecutive outpatients who were referred for PSG evaluation for suspected sleep-disordered breathing (SDB) by their primary care physicians. If the same patient underwent more than one test, only the test that had the highest total sleep time was retained. All subjects were grouped into two different and independent samples, using a randomization process as prescribed by the SPSS statistical package (version 17.0; SPSS; Chicago, Illinois, United States), in which all patients enrolled in the study were divided equally into two cohorts (50% of the patients were allocated to the derivation cohort and the remaining 50% were allocated to the validation cohort). Our study protocol (#666.608) was approved by the Institutional Review Board of Federal University of Rio de Janeiro and waived the patient consent requirement.

For both groups, inclusion criteria consisted of age 18 years or older and suspected OSA, whereas they were excluded for any of the following reasons: previously diagnosed OSA, use of portable or split-night studies, incomplete clinical data, and technically inadequate PSG. On the evening of the PSG, clinical data were collected in all patients: sex, age, BMI, NC, self-reported comorbidities (hypertension, diabetes mellitus, and smoking), and Epworth Sleepiness Scale (ESS),[17] which is a well-validated 8-item questionnaire that measures subjective sleepiness, with a score of 10 points or higher considered indicative of excessive daytime sleepiness. Patients were measured for weight and height, and the BMI was calculated by dividing the weight in kilograms by the square of the height in meters (kg/m^2). The NC was systematically measured, using a tape measure, in centimeters, as follows: all subjects were asked to stand erect with their head positioned in the Frankfort horizontal plane. The superior border of a tape measure was placed just below the laryngeal prominence and applied perpendicular to the long axis of the neck.[18]

Aiming to compare with other validated models, data from the STOP-BANG questionnaire[13] and NoSAS score[16] were also collected by trained sleep technicians. The STOP-BANG[13] consists of 8 yes-or-no questions: loud snoring, tiredness, observed apnea, hypertension, BMI > 35 kg/m^2, age older than 50 years, NC > 40 cm, and male sex. A score of 3 or higher is considered as high risk for presence of OSA. The NoSAS score[16] allocates 4 points for having a NC > 40 cm, 3 points for having a BMI of 25 kg/m^2 to less than 30 kg/m^2 or 5 points for having a BMI ≥ 30 kg/m^2, 2 points for snoring, 4 points for age older than 55 years, and 2 points for being male. Using NoSAS, a cutoff of ≥ 8 points identifies subjects at risk of clinically significant SDB.[16]

Sleep Studies

All Brazilian studies were conducted in a single sleep center with two different units: one in Niteroi City and one in Rio de Janeiro City. All patients underwent an attended, in-laboratory PSG (EMBLA S7000, Embla Systems, Inc., Broomfield, Colorado, United States), consisting of continuous monitoring of electroencephalography, electrooculography, electromyography (chin and legs), electrocardiography, airflow (nasal pressure), thoracic and abdominal impedance belts, oxygen saturation (SpO_2), microphone for snoring, and sensors for body position. PSG records were scored manually and were interpreted in a blinded way by two board-certified sleep physicians in accordance with existing guidelines.[19] Apneas were classified with a drop ≥ 90% of baseline in airflow lasting at least 10 seconds, whereas hypopneas were classified as follows: a drop ≥ 30% of preevent during ≥ 10 seconds associated with ≥ 3% oxygen desaturation or an arousal.[19] Diagnosis of OSA was based on an apnea-hypopnea index (AHI) ≥ 5 events/h and its severity was classified as follows: ≥ 5 events/h as any OSA (OSA-5), ≥ 15 events/h as moderate/severe OSA (OSA-15), and ≥ 30 events/h as severe OSA (OSA-30).

Modeling

We aimed to elaborate a very practical model that contains only continuous and numerical variables, and as such three parameters, previously recognized as predictors of OSA, were chosen for analysis: NC, age, and BMI. First, these variables were evaluated through linear regression in order to verify the occurrence or not of multicollinearity. Multicollinearity is a linear association between two or more explanatory variables being detected by tolerance and variance inflation factor (VIF).[20] The VIF is simply the reciprocal of the tolerance level (1/tolerance), being that the VIFs measure the inflation in the variances of the parameter estimates due to collinearities that exist among the predictors.[20]

After excluding multicollinearity, these three parameters were evaluated in their continuous form, through binary logistic regression at three different AHI thresholds (5, 15, and 30 events/h), with comparisons being conducted through the regression coefficient (β). To further simplify the model, the two main parameters were retained. For the elaboration of the scoring system, these two parameters were grouped based on three cutoff points obtained by median (50% percentile) and interquartile range (IQR: 25% and 75% percentiles) and then evaluated with binary logistic regression. Assignment of points to parameters was calculated as follows: the β coefficient of each variable was divided by the lowest β-coefficient value and rounded to the nearest whole integer. Subsequently, all points accrued were summed to create the scoring index.

Statistical Analysis

Data analysis was conducted using SPSS for Windows (version 17.0; SPSS; Chicago, Illinois, United States). Results are shown as median and IQR for continuous variables and as frequency with percentage for categorical variables. Groups were compared using the chi-square test for categorical variables, whereas all numeric variables were evaluated with the nonparametric Mann-Whitney *U* test. Correlation was evaluated by Spearman correlation coefficient (r_s). Through linear regression, absence of multicollinearity was assumed if VIF < 5.0. Parameters predictive of OSA (age, NC, and BMI) were entered into a logistic regression analysis in parallel, with the Wald test being used to explore if explanatory variables in a model were significant. A scoring system was derived using weightings from β-regression coefficients.

To examine the apparent performance (internal validity) of our developed model, discrimination, calibration, and overall performance were calculated in both cohorts evaluated. Discrimination, the ability of a scoring system to distinguish between patients with and without different outcomes, was estimated from the area under the curve (AUC).[21] The AUC may theoretically range from 0.5 (discrimination equivalent to that of chance) to 1.0 (perfect discrimination).[21] Calibration refers to the agreement between observed outcomes and predictions; being that it was assessed by Hosmer-Lemeshow chi-square test ($P < .05$ indicates poor calibration).[21] Overall performance (how well the model predicts the likelihood of an outcome in an individual patient) was assessed using the Nagelkerke R^2, which ranges from 0 to 1.[21] As the Hosmer-Lemeshow chi-square test is sensitive to sample size, we chose smaller subsets of randomly selected patients (n = 1,000) to evaluate the model calibration in both cohorts evaluated. After completion of the model development, the cutoff point determined to identify patients at risk for SDB was chosen to achieve a high sensitivity, while preserving a moderate specificity. Using the 2 × 2 contingency tables, the following parameters were calculated: sensitivity, specificity, PPV, NPV, accuracy, likelihood ratios, and odds ratio (OR) with their respective 95% confidence interval (CI). The receiver operating characteristic (ROC) curves and AUC were assessed at three AHI thresholds (5, 15, and 30 events/h). The AUCs obtained by all screening models were compared using prior algorithm.[22] Posttest probability of each score was calculated by logistic regression. A two-tailed value of $P < .05$ was considered statistically significant.

RESULTS

Of a total of 4,476 subjects referred for diagnostic PSG, 404 patients (9.0%) were subsequently excluded based on exclusion criteria. The exclusions consisted of 260 with incomplete clinical data, 76 tested with portable or split-night studies, 54 with technically inadequate PSG, and 14 with a previous diagnosis of OSA. Thus, 4,072 subjects were randomly allocated into two independent cohorts: one for derivation (n = 2,037) and one for validation (n = 2,035). Based on **Table 1**, no clinical or PSG parameter was statistically different. Within the derivation cohort, median age was 45.0 years (IQR: 35.0–55.0) and 55.9% were male, whereas in the derivation cohort, median age was 44.0 years (IQR: 34.0–55.0) and 53.2% were male. Similarly, the prevalence of OSA-5, OSA-15, and OSA-30 was not statistically different between the derivation cohort and validation cohort (77.9% versus 76.4%, 55.2% versus 54.7%, and 34.5% and 35.8%; respectively). In the derivation cohort, prevalence of OSA-5, OSA-15, and OSA-30 was higher in men than women: 86.9% versus 66.5%, 67.6% versus 39.5%, and 46.7% versus 19.0%; respectively; all with $P < .001$. Similarly, in the validation cohort, prevalence of OSA-5, OSA-15, and OSA-30 were also higher in men than women: 87.2% versus 64.1%, 69.3% versus 38.1%, and 48.9% versus 20.8%; respectively; all with $P < .001$, suggesting sex influences leading to higher prevalence of severe forms of OSA. Men had a higher median NC than women in the derivation cohort (42.0 cm [IQR: 40.0–46.0] versus 38.0 cm [IQR: 35.0–40.0]), and in the validation cohort (43.0 cm [IQR: 40.0–46.0] versus 38.0 cm [IQR: 35.0–40.0]), both with $P < .001$. Men had a lower median age than women in the derivation cohort (42.0 years [IQR: 33.0–53.0] versus 47.0 years [IQR: 35.0–56.0]) and in the validation cohort (41.0 years [IQR: 33.0–53.0] versus 47.0 years [IQR: 36.0–58.0]); both with $P < .001$.

The No-Apnea Development

After excluding multicollinearity among the three parameters of interest (NC with VIF = 1.484, age with VIF = 1.016, and BMI with VIF = 1.503), these variables were evaluated through binary logistic regression (**Table 2**). As can be seen in **Table 2**, for all AHI thresholds (5, 15 or 30 events/h), when comparing the three clinical parameters in their continuous form, we found that the magnitude of the regression coefficient was always higher for NC, followed by age and BMI. Moreover, BMI did not emerge as an independent variable for screening of OSA-5 with OR: 1.014 (95% CI: 0.996–1.033; $P = .119$). Therefore, BMI was then excluded from the model for three reasons: (1) in the binary logistic regression, it was the third-ranked parameter for all AHI thresholds; (2) it did not emerge as an independent parameter for screening of OSA-5; and (3) reducing the model to two variables obviously further simplified the tool.

Table 3 shows the binary logistic regression with the two parameters chosen (NC and age) categorized from the 25%, 50% and 75% percentiles: (1) for NC: 37 cm, 40 cm, and 43 cm; and (2) for age: 35 years, 45 years, and 55 years; respectively.

Table 1—General and sleep characteristics for both cohorts.

Parameter	Derivation Cohort (n = 2,037)	Validation Cohort (n = 2,035)	*P*
Clinical data			
Male, n (%)	1,138 (55.9)	1,083 (53.2)	.095
Caucasian, n (%)	1,636 (80.3)	1,645 (80.8)	.864
Age, years	45.0 (35.0–55.0)	44.0 (34.0–55.0)	.975
BMI, kg/m^2	32.2 (26.7–38.8)	32.5 (26.7–38.9)	.552
NC, cm	40.0 (37.0–43.0)	40.0 (37.0–44.0)	.660
ESS score ≥ 10, n (%)	1,017 (50.0)	1,063 (52.3)	.167
Current smokers, n (%)	202 (9.9)	202 (9.9)	> .999
Hypertension, n (%)	819 (40.2)	790 (38.8)	.370
Diabetes mellitus, n (%)	224 (11.0)	250 (12.3)	.204
PSG data			
Total sleep time, min	352.1 (310.4–393.0)	358.5 (307.8–395.5)	.064
Sleep efficiency, %	82.2 (71.7–90.5)	83.0 (71.4–90.6)	.198
Sleep latency, min	22.4 (9.0–48.3)	22.4 (9.8–47.0)	.950
REM latency, min	126.0 (86.5–188.2)	125.5 (85.2–187.0)	.847
Stage N1, %	3.0 (1.0–7.0)	3.0 (1.0–7.0)	.863
Stage N2, %	65.5 (58.0–74.0)	65.0 (57.1–74.1)	.354
Stage N3, %	12.2 (4.2–18.8)	12.0 (4.7–18.5)	.730
Stage R, %	16.5 (11.2–21.5)	16.8 (11.2–22.0)	.315
Arousal index, events/h	21.5 (11.4–42.4)	20.8 (10.6–42.3)	.442
AHI, events/h	16.9 (6.2–41.3)	17.3 (5.5–40.8)	.531
Awake SpO_2, %	95.8 (94.5–96.9)	95.8 (94.5–97.0)	.488
Mean SpO_2, %	94.2 (92.0–95.9)	94.3 (92.0–95.9)	.864
Nadir SpO_2, %	84.0 (77.0–89.0)	84.0 (77.0–89.0)	.222
AHI, events/h			
≥ 5, n (%)	1,587 (77.9)	1,554 (76.4)	.247
≥ 15, n (%)	1,124 (55.2)	1,114 (54.7)	.801
≥ 30, n (%)	702 (34.5)	728 (35.8)	.393

Data are presented as median (interquartile range) or n (%). AHI = apnea-hypopnea index, BMI = body-mass index, ESS = Epworth sleepiness scale, NC = neck circumference, PSG = polysomnography, REM = rapid eye movement, SpO_2 = oxygen saturation.

Table 2—Binary logistic regression of the continuous predictors evaluated according to AHI thresholds (derivation cohort: n = 2,037).

	β	SE	Wald	*df*	*P*	OR (95% CI)
AHI ≥ 5 events/h						
NC, cm	0.235	0.018	174.078	1	< .001	1.265 (1.222–1.311)
Age, years	0.047	0.005	100.267	1	< .001	1.048 (1.039–1.058)
BMI, kg/m^2	0.014	0.009	2.436	1	.119	1.014 (0.996–1.033)
AHI ≥ 15 events/h						
NC, cm	0.204	0.015	196.975	1	< .001	1.226 (1.191–1.261)
Age, years	0.033	0.004	74.138	1	< .001	1.033 (1.026–1.041)
BMI, kg/m^2	0.024	0.008	10.227	1	.001	1.025 (1.010–1.040)
AHI ≥ 30 events/h						
NC, cm	0.215	0.015	207.453	1	< .001	1.240 (1.204–1.277)
Age, years	0.025	0.004	39.967	1	< .001	1.025 (1.017–1.033)
BMI, kg/m^2	0.021	0.008	6.746	1	.009	1.021 (1.005–1.037)

The *P* value was obtained from the Wald test. All parameters were entered into logistic regression in parallel. AHI = apnea-hypopnea index, β = regression coefficient, BMI = body mass index, CI = confidence interval, *df* = degrees of freedom for the Wald test, NC = neck circumference, OR = odds ratio, SE = standard error.

Table 3—Binary logistic regression of the categorized predictors according to AHI ≥ 5 events/h (derivation cohort: n = 2,037).

	β	Points*	SE	Wald	*df*	*P*	OR (95% CI)
Neck circumference, cm							
< 37.0	–	0	–	221.440	3	< .001	–
37.0–39.9	0.732	+ 1	0.153	22.941	1	< .001	2.080 (1.541–2.807)
40.0–42.9	1.422	+ 3	0.157	82.538	1	< .001	4.145 (3.050–5.633)
≥ 43.0	3.028	+ 6	0.215	198.101	1	< .001	20.660 (13.552–31.497)
Age, years							
< 35	–	0	–	100.516	3	< .001	–
35–44	0.492	+ 1	0.154	10.209	1	.001	1.636 (1.210–2.213)
45–54	0.981	+ 2	0.163	36.120	1	< .001	2.668 (1.937–3.675)
≥ 55	1.681	+ 3	0.175	91.951	1	< .001	5.369 (3.808–7.569)

The *P* value was obtained from the Wald test. * = points assigned to No-Apnea from the regression coefficient. AHI = apnea-hypopnea index, β = regression coefficient, CI = confidence interval, *df* = degrees of freedom for the Wald test, OR = odds ratio, SE = standard error.

Then, each variable was grouped and scored, according to β-coefficient, as follows: the NC (in cm) was scored in three different values: 1 (37.0–39.9), 3 (40.0–42.9), and 6 (≥ 43.0), whereas the age (in years) was scored in three different values: 1 (35–44), 2 (45–54), and 3 (≥ 55). The points for each variable were added, totaling a final score of 0–9 points and this mnemonic tool was termed "No-Apnea" (**Table 4**). The most frequent No-Apnea score was 3 points (n = 352), followed by 6 points (n = 314) and 4 points (n = 263). Corresponding to the increase in No-Apnea scores (from 0 to 9 points), there was a linear increase in the prevalence of OSA-5 (from 33.0% to 95.6%), OSA-15 (from 13.6% to 85.3%), and OSA-30 (from 3.9% to 68.4%); all with $P < .001$.

Table 4—No-Apnea scoring system.

Parameter	Points
Neck circumference, cm	
< 37.0	0
37.0–39.9	+ 1
40.0–42.9	+ 3
≥ 43.0	+ 6
Age, years	
< 35	0
35–44	+ 1
45–54	+ 2
≥ 55	+ 3

The points for each variable are added, totaling a final score of 0–9 points.

The No-Apnea Performance

In the derivation cohort, for screening of OSA-5 (the AHI threshold chosen to score the No-Apnea parameters), the developed model showed the following characteristics: (1) discriminatory power based on ROC curves with AUC: 0.784 (95% CI: 0.761–0.808); (2) calibration with Hosmer-Lemeshow chi-square test: 10.270 ($P = .247$); and (3) overall performance with Nagelkerke R^2: 0.265. Moreover, for screening of OSA-15 and OSA-30, the No-Apnea showed the following characteristics: (1) discriminatory power with AUCs: 0.758 (95% CI: 0.737–0.779) and 0.754 (95% CI: 0.733–0.776), respectively; (2) calibration with Hosmer-Lemeshow chi-square test: 11.591 ($P = .170$) and 10.046 ($P = .262$), respectively; and (3) overall performance with Nagelkerke R^2: 0.259 and 0.248, respectively.

In the validation cohort, for screening of OSA-5, OSA-15, and OSA-30, the No-Apnea showed the following characteristics: (1) discriminatory power with AUCs: 0.781 (95% CI: 0.757–0.805), 0.752 (95% CI: 0.731–0.773), and 0.752 (95% CI: 0.730–0.773), respectively; (2) calibration with Hosmer-Lemeshow chi-square test: 10.976 ($P = .203$), 13.647 ($P = .091$), and 5.498 ($P = .703$), respectively; and (3) overall performance with Nagelkerke R^2: 0.259, 0.243, and 0.234, respectively.

The No-Apnea Predictive Parameters (Derivation Cohort)

Predictive performance of the No-Apnea is shown in **Table 5** (derivation cohort; n = 2,037). We used a cutoff of ≥ 3 to classify patients at high risk (75.9%) versus at low risk (24.1%) of having OSA-5, OSA-15, and OSA-30. The accuracies obtained were of 78.1%, 68.8%, and 54.4%, respectively for OSA-5, OSA-15, and OSA-30. Using a cutoff of ≥ 3, the posttest probabilities for OSA-5, OSA-15, and OSA-30 were 86.8%, 65.8%, and 42.6%, respectively. In addition, the posttest probabilities for OSA-5, OSA-15, and OSA-30 increased proportionally with the increase in the No-Apnea scores (from 0 to 9 points; data not shown).

Table 6 was created aiming to compare our model with two previously reported screening tools: STOP-BANG and NoSAS. For screening of OSA-5, OSA-15, and OSA-30, No-Apnea model showed the following parameters: sensitivity ranged from 84.7% to 94.0%, specificity ranged from 54.9% to 33.6%, whereas the accuracy ranged from 78.1% to 54.4%, respectively. For OSA-5 diagnosis, STOP-BANG showed the higher sensitivity (88.8%), whereas the higher specificity was obtained with NoSAS (68.7%). For OSA-15 diagnosis, NoSAS showed the higher specificity (57.6%), whereas the higher sensitivity was obtained with STOP-BANG (92.5%). For OSA-30 diagnosis, NoSAS showed the higher specificity (49.2%), whereas the higher sensitivity was obtained with STOP-BANG (95.7%). Based on AUCs, No-Apnea discrimination did not show statistically significant differences compared to the STOP-BANG for screening of OSA-5

Table 5—Predictive parameters of No-Apnea (derivation cohort: n = 2,037).

	No-Apnea Scores		
	≥ 2 versus < 2	≥ 3 versus < 3	≥ 4 versus < 4
AHI ≥ 5 events/h			
Sensitivity	92.1 (91.2–93.1)	84.7 (83.6–85.8)	68.1 (67.0–69.1)
Specificity	38.2 (34.8–41.5)	54.9 (51.0–58.6)	74.7 (70.8–78.2)
PPV	84.0 (83.1–84.9)	86.9 (85.8–88.0)	90.5 (89.0–91.8)
NPV	57.9 (52.8–62.9)	50.4 (46.9–53.9)	39.9 (37.8–41.8)
Accuracy	80.2 (78.7–81.7)	78.1 (76.4–79.8)	69.6 (67.9–71.1)
LR +	1.49 (1.39–1.59)	1.87 (1.70–2.07)	2.68 (2.29–3.17)
LR −	0.20 (0.16–0.25)	0.27 (0.24–0.32)	0.42 (0.39–0.46)
Odds ratio	7.23 (5.51–9.50)	6.73 (5.30–8.53)	6.29 (4.93–8.04)
Posttest probability (%)	84.0	86.8	90.4
AHI ≥ 15 events/h			
Sensitivity	95.1 (93.9–96.2)	90.6 (89.0–92.0)	76.2 (74.4–78.0)
Specificity	26.5 (25.0–27.8)	42.1 (40.2–43.8)	63.0 (60.7–65.2)
PPV	61.4 (60.6–62.1)	65.8 (64.7–66.8)	71.7 (70.0–73.4)
NPV	81.5 (76.8–85.5)	78.4 (74.8–81.6)	68.3 (65.8–70.7)
Accuracy	64.4 (63.0–65.5)	68.8 (67.1–70.4)	70.3 (68.3–72.3)
LR +	1.29 (1.25–1.33)	1.56 (1.48–1.63)	2.06 (1.89–2.24)
LR −	0.18 (0.13–0.24)	0.22 (0.18–0.27)	0.37 (0.33–0.42)
Odds ratio	7.01 (5.09–9.66)	6.97 (5.45–8.92)	5.46 (4.48–6.64)
Posttest probability (%)	61.4	65.8	71.7
AHI ≥ 30 events/h			
Sensitivity	97.3 (95.8–98.3)	94.0 (92.1–95.5)	81.9 (79.3–84.4)
Specificity	20.8 (20.1–21.4)	33.6 (32.6–34.4)	53.6 (52.2–54.8)
PPV	39.3 (38.7–39.7)	42.7 (41.8–43.4)	48.1 (46.6–49.6)
NPV	93.6 (90.1–96.0)	91.4 (88.7–93.6)	84.9 (82.7–87.0)
Accuracy	47.2 (46.2–47.9)	54.4 (53.1–55.4)	63.3 (61.5–65.0)
LR +	1.22 (1.19–1.25)	1.41 (1.36–1.45)	1.76 (1.65–1.86)
LR −	0.13 (0.08–0.20)	0.17 (0.13–0.24)	0.33 (0.28–0.39)
Odds ratio	9.45 (5.76–15.67)	7.93 (5.63–11.22)	5.22 (4.16–6.54)
Posttest probability (%)	39.1	42.6	48.1

No-Apnea scoring system is a 2-item model: neck circumference is scored as follows: 37.0–39.9 cm (1 point), 40.0–42.9 cm (3 points), and ≥ 43.0 cm (6 points), whereas age is scored as follows: 35–44 years (1 point), 45–54 years (2 points), ≥ 55 years (3 points); totaling a score of 0–9 points. Data are presented as estimates (95% confidence intervals) unless otherwise stated. AHI = apnea-hypopnea index, LR = likelihood ratio, NPV = negative predictive value, PPV = positive predictive value.

(*P* = .645), OSA-15 (*P* = .946), and OSA-30 (*P* = .589). Similarly, No-Apnea discrimination did not significantly differ from that of NoSAS for screening of OSA-5 (*P* = .555), OSA-15 (*P* = .946), and OSA-30 (*P* = .858). Furthermore, STOP-BANG discrimination was similar to NoSAS for diagnosis of OSA-5, OSA-15, and OSA-30: *P* = .896, *P* = .892, and *P* = .473; respectively. All models tested (No-Apnea, STOP-BANG, and NoSAS) were correlated with AHI (r_s = 0.530, r_s = 0.545, and r_s = 0.529; respectively; all with *P* < .001). In addition, ESS has not proved useful as a screening tool for OSA-5, OSA-15, and OSA-30: AUC: 0.573 (95% CI: 0.543–0.603), AUC: 0.559 (95% CI: 0.534–0.584), and AUC: 0.591 (95% CI: 0.565–0.617); respectively.

The No-Apnea Predictive Parameters (Validation Cohort)

Based on AUCs summarized in the **Table 7** (validation cohort; n = 2,035), No-Apnea discrimination did not show statistically significant differences compared to the STOP-BANG for screening of OSA-5 (*P* = .232), OSA-15 (*P* = .087), and OSA-30 (*P* = .074). Similarly, No-Apnea discrimination did not significantly differ from that of NoSAS for screening of OSA-5 (*P* = .957), OSA-15 (*P* = .788), and OSA-30 (*P* > .999). In addition, STOP-BANG discrimination was similar to NoSAS for diagnosis of OSA-5, OSA-15, and OSA-30: *P* = .212, *P* = .085, and *P* = .074; respectively. For screening of OSA-5, OSA-15, and of OSA-30, No-Apnea model showed the following parameters: sensitivity ranged from 83.1% to 91.5%, specificity ranged from 58.2% to 36.8%, whereas the accuracy ranged from 77.2% to 56.4%, respectively. All models tested (No-Apnea, STOP-BANG, and NoSAS) were correlated with AHI (r_s = 0.528, r_s = 0.584, and r_s = 0.534; respectively; all with *P* < .001). All AUCs obtained by the three models are shown in **Figure 1**.

Table 6—Comparing predictive performances of the No-Apnea, STOP-BANG, and NoSAS (derivation cohort: n = 2,037).

	Screening Tools		
	No-Apnea	**STOP-BANG**	**NoSAS**
AHI ≥ 5 events/h (prevalence: 77.9%)			
Sensitivity	84.7 (83.6–85.8)	88.8 (87.7–89.8)	71.6 (70.5–72.6)
Specificity	54.9 (51.0–58.6)	47.1 (43.4–50.7)	68.7 (64.7–72.4)
PPV	86.9 (85.8–88.0)	85.5 (84.5–86.5)	89.0 (87.6–90.3)
NPV	50.4 (46.9–53.9)	54.4 (50.1–58.5)	40.7 (38.3–42.9)
Accuracy	78.1 (76.4–79.8)	79.6 (78.0–81.2)	70.9 (69.2–72.6)
LR +	1.87 (1.70–2.07)	1.67 (1.55–1.82)	2.28 (1.99–2.63)
LR −	0.27 (0.24–0.32)	0.23 (0.20–0.28)	0.41 (0.37–0.45)
Odds ratio	6.73 (5.30–8.53)	7.05 (5.49–9.04)	5.52 (4.37–6.97)
AUC	0.784 (0.761–0.808)	0.777 (0.752–0.801)	0.775 (0.752–0.799)
AHI ≥ 15 events/h (prevalence: 55.2%)			
Sensitivity	90.6 (89.0–92.0)	92.5 (91.1–93.8)	79.2 (77.3–80.9)
Specificity	42.1 (40.2–43.8)	33.5 (31.8–35.1)	57.6 (55.4–59.8)
PPV	65.8 (64.7–66.8)	63.1 (62.2–64.0)	69.7 (68.1–71.2)
NPV	78.4 (74.8–81.6)	78.5 (74.4–82.1)	69.2 (66.5–71.8)
Accuracy	68.8 (67.1–70.4)	66.1 (64.5–67.5)	69.5 (67.5–71.5)
LR +	1.56 (1.48–1.63)	1.39 (1.33–1.44)	1.86 (1.73–2.01)
LR −	0.22 (0.18–0.27)	0.22 (0.17–0.28)	0.36 (0.31–0.40)
Odds ratio	6.97 (5.45–8.92)	6.24 (4.76–8.17)	5.17 (4.23–6.31)
AUC	0.758 (0.737–0.779)	0.759 (0.738–0.779)	0.757 (0.736–0.778)
AHI ≥ 30 events/h (prevalence: 34.5%)			
Sensitivity	94.0 (92.1–95.5)	95.7 (94.0–97.0)	85.3 (82.8–87.6)
Specificity	33.6 (32.6–34.4)	27.0 (26.1–27.6)	49.2 (47.9–50.4)
PPV	42.7 (41.8–43.4)	40.8 (40.1–41.3)	46.9 (45.5–48.2)
NPV	91.4 (88.7–93.6)	92.3 (89.2–94.6)	86.4 (84.1–88.5)
Accuracy	54.4 (53.1–55.4)	50.7 (49.5–51.5)	61.7 (59.9–63.2)
LR +	1.41 (1.36–1.45)	1.31 (1.27–1.34)	1.68 (1.58–1.76)
LR −	0.17 (0.13–0.24)	0.15 (0.10–0.22)	0.29 (0.24–0.35)
Odds ratio	7.93 (5.63–11.22)	8.27 (5.54–12.40)	5.63 (4.42–7.18)
AUC	0.754 (0.733–0.776)	0.763 (0.742–0.784)	0.751 (0.730–0.773)

Data are presented as estimates (95% confidence intervals) unless otherwise stated. No-Apnea scoring system is a 2-item model: NC is scored as follows: 37.0–39.9 cm (1 point), 40.0–42.9 cm (3 points), and ≥ 43.0 cm (6 points), whereas age is scored as follows: 35–44 years (1 point), 45–54 years (2 points), ≥ 55 years (3 points); totaling a score of 0–9 points (score ≥ 3 was considered as high-risk for presence of any OSA, moderate/severe OSA, and severe OSA). STOP-BANG questionnaire is an 8-item model (1 point for each positive answer): loud snoring, tiredness, observed apnea, hypertension, BMI > 35 kg/m^2, age > 50 years, NC > 40 cm, and male sex; totaling a score of 0–8 points (score ≥ 3 was considered as high-risk for presence of any OSA, moderate/severe OSA, and severe OSA). NoSAS score is a 5-item model: NC > 40 cm (4 points), BMI 25.0–29.9 kg/m^2 (3 points), BMI ≥ 30.0 kg/m^2 (5 points), snoring (2 points), age > 55 years (4 points), male sex (2 points); totaling a score of 0–17 points (score ≥ 8 was considered as high risk for presence of any OSA, moderate/severe OSA, and severe OSA). AHI = apnea-hypopnea index, AUC = area under the curve, BMI = body mass index, LR = likelihood ratio, NC = neck circumference, NPV = negative predictive value, OSA = obstructive sleep apnea, PPV = positive predictive value.

DISCUSSION

Our findings show that, in a sleep-laboratory setting containing a sample of subjects referred for evaluation of suspected OSA, an extremely simple tool, No-Apnea, designed with only two numeric and objectively acquired variables (NC and age) exhibits excellent and reproducible performance when screening for OSA at any level of severity. In both derivation and validation cohorts, the No-Apnea discrimination (a 2-item tool) did not significantly differ from that of STOP-BANG (an 8-item tool) and NoSAS (a 5-item tool). As expected, both cohorts presented a high prevalence of OSA, an anticipated finding because these were clinically referred sleep-laboratory patients, a population known to have a high prevalence of OSA. According to previous studies,[8,10,23–25] our findings showed that males had higher NC than females, whereas females were older when compared with males. Similarly, we also observed that men had a higher rate of OSA compared to women.[3–5,8,10,23–25]

The cutoff used to classify patients with high pretest probability of OSA was 3 points for OSA-5, OSA-15, and OSA-30. The inclusion of this single cutoff was chosen to obtain a high sensitivity with consequent moderate specificity. Sensitivity and specificity of a screening model are usually inversely related, and the high sensitivity often comes at the expense of

Table 7—Comparing predictive performances of the No-Apnea, STOP-BANG, and NoSAS (validation cohort: n = 2,035).

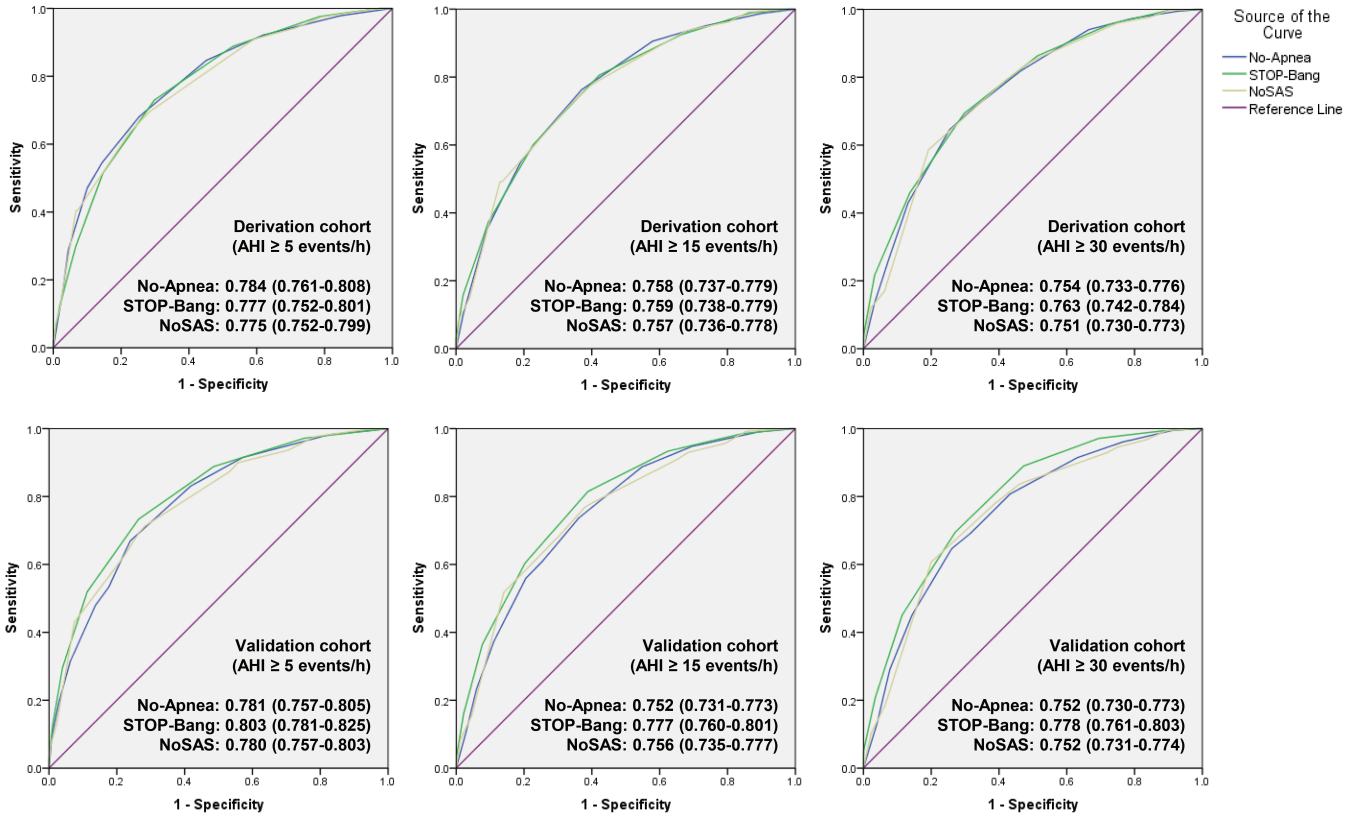
	Screening Tools		
	No-Apnea	**STOP-BANG**	**NoSAS**
AHI ≥ 5 events/h (prevalence: 76.4%)			
Sensitivity	83.1 (81.9–84.2)	88.9 (87.8–89.9)	71.3 (70.1–72.4)
Specificity	58.2 (54.5–61.8)	51.4 (47.8–54.7)	71.7 (68.0–75.2)
PPV	86.5 (85.3–87.7)	85.5 (84.5–86.5)	89.1 (87.6–90.4)
NPV	51.6 (48.3–54.7)	58.8 (54.8–62.7)	43.6 (41.3–45.7)
Accuracy	77.2 (75.4–78.9)	80.0 (78.3–81.6)	71.4 (69.6–73.1)
LR +	1.98 (1.80–2.20)	1.82 (1.68–1.98)	2.52 (2.19–2.92)
LR −	0.29 (0.25–0.33)	0.21 (0.18–0.25)	0.40 (0.36–0.43)
Odds ratio	6.83 (5.42–8.61)	8.42 (6.59–10.77)	6.30 (4.99–7.96)
AUC	0.781 (0.757–0.805)	0.803 (0.781–0.825)	0.780 (0.757–0.803)
AHI ≥ 15 events/h (prevalence: 54.7%)			
Sensitivity	88.7 (87.1–90.2)	93.4 (92.0–94.6)	78.3 (76.4–80.1)
Specificity	45.3 (43.3–47.1)	37.6 (35.9–39.1)	59.6 (57.4–61.8)
PPV	66.2 (65.0–67.3)	64.4 (63.4–65.2)	70.1 (68.4–71.7)
NPV	76.8 (73.5–79.9)	82.4 (78.7–85.6)	69.4 (66.8–71.9)
Accuracy	69.0 (67.3–70.7)	68.1 (66.6–69.5)	69.8 (67.8–71.8)
LR +	1.62 (1.53–1.70)	1.49 (1.43–1.55)	1.93 (1.79–2.09)
LR −	0.25 (0.20–0.29)	0.17 (0.13–0.22)	0.36 (0.32–0.41)
Odds ratio	6.48 (5.14–8.19)	8.45 (6.39–11.19)	5.31 (4.36–6.48)
AUC	0.752 (0.731–0.773)	0.777 (0.760–0.801)	0.756 (0.735–0.777)
AHI ≥ 30 events/h (prevalence: 35.8%)			
Sensitivity	91.5 (89.4–93.3)	97.1 (95.6–98.1)	84.5 (82.0–86.7)
Specificity	36.8 (35.6–37.8)	30.5 (29.7–31.1)	51.9 (50.5–53.1)
PPV	44.6 (43.6–45.5)	43.8 (43.1–44.2)	49.4 (48.0–50.8)
NPV	88.6 (85.8–91.0)	95.0 (92.4–96.8)	85.7 (83.4–87.8)
Accuracy	56.4 (54.9–57.6)	54.3 (53.3–55.1)	63.5 (61.8–65.2)
LR +	1.44 (1.38–1.49)	1.39 (1.36–1.42)	1.75 (1.65–1.85)
LR −	0.23 (0.17–0.29)	0.09 (0.06–0.14)	0.29 (0.24–0.35)
Odds ratio	6.25 (4.67–8.39)	14.79 (9.26–23.84)	5.86 (4.64–7.41)
AUC	0.752 (0.730–0.773)	0.778 (0.761–0.803)	0.752 (0.731–0.774)

Data are presented as estimates (95% confidence intervals) unless otherwise stated. No-Apnea scoring system is a 2-item model: NC is scored as follows: 37.0–39.9 cm (1 point), 40.0–42.9 cm (3 points), and ≥ 43.0 cm (6 points); whereas age is scored as follows: 35–44 years (1 point), 45–54 years (2 points), ≥ 55 years (3 points); totaling a score of 0–9 points (score ≥ 3 was considered as high-risk for presence of any OSA, moderate/severe OSA, and severe OSA). STOP-BANG questionnaire is an 8-item model (1 point for each positive answer): loud snoring, tiredness, observed apnea, hypertension, BMI > 35 kg/m^2, age > 50 years, NC > 40 cm, and male sex; totaling a score of 0–8 points (score ≥ 3 was considered as high-risk for presence of any OSA, moderate/severe OSA, and severe OSA). NoSAS score is a 5-item model: NC > 40 cm (4 points), BMI 25.0–29.9 kg/m^2 (3 points), BMI ≥ 30.0 kg/m^2 (5 points), snoring (2 points), age > 55 years (4 points), male sex (2 points); totaling a score of 0–17 points (score ≥ 8 was considered as high risk for presence of any OSA, moderate/severe OSA, and severe OSA). AHI = apnea-hypopnea index, AUC = area under the curve, BMI = body mass index, LR = likelihood ratio, NC = neck circumference, NPV = negative predictive value, OSA = obstructive sleep apnea, PPV = positive predictive value.

specificity. For a disease such as OSA, it is possibly more important that a screening test has a high sensitivity, and does not miss patients with OSA, rather than a high specificity, especially in a population with high pretest probability of disease.[10] However, this strategy may not be unanimous. A previous study[26] reports on the strategy of using different cutoffs to rule-in or rule-out OSA. In a sleep-laboratory setting, where the main objective is to identify subjects with more severe forms of OSA and requiring treatment with continuous positive airway pressure (CPAP), a higher cutoff for a given model may be preferred, whereas in a primary care setting where the priority is not to miss any disease, a lower cutoff may be more appropriate.[26] Similar to prior studies on STOP-BANG,[27–29] as the No-Apnea score increased, the posttest probability of having OSA-5, OSA-15, and OSA-30 also increased.

OSA is a very prevalent and often underdiagnosed disease.[30,31] In addition, symptoms suggestive of OSA, despite being common, are not consistently investigated during routine clinical visits.[32] Accordingly, screening tools can be used to identify patients at high risk for SDB, thus prioritizing the use of portable methods in areas with limited resources. However, the performance of an OSA questionnaire may have considerable variability according to the patient population and AHI thresholds employed.[10,33] The Berlin questionnaire, for

Figure 1—Receiver operating characteristic curves showing the discrimination of No-Apnea, STOP-BANG, and NoSAS.



Values shown as area under the curve and 95% confidence interval. Top panels: derivation cohort (n = 2,037). Bottom panels: validation cohort (n = 2,035). OSA severity was classified based on AHI as follows: ≥ 5 events/h as any OSA, ≥ 15 events/h as moderate/severe OSA, and ≥ 30 events/h as severe OSA. AHI = apnea-hypopnea index, OSA = obstructive sleep apnea.

example, possibly has a better performance in primary care settings than sleep laboratory settings.[26,34] In a sleep clinic population, the STOP-BANG showed sensitivity of 90%, 94%, and 96% to detect OSA-5, OSA-15, and OSA-30, respectively; however, specificity was relatively low (49%, 34%, and 25%; respectively).[35] Conversely, STOP-BANG alone was insufficient to confirm the occurrence of significant OSA in military veterans undergoing unattended sleep studies, mainly OSA-15, in whom a score of 3 showed a high sensitivity (99.1%), but also a very low specificity (4.9%).[36]

The STOP-BANG is an instrument with high sensitivity that increases in parallel with increasing AHI thresholds (from 5 to 30 events/h).[13] Conversely, it exhibits moderate to low specificity based on the AHI thresholds used herein, such that decreases in specificity result in a large number of false-positive results, thereby reducing accuracy, especially in the more severe forms of OSA. This issue is critical, particularly in sleep laboratories targeting the choice of portable diagnostic methods or when predicting which subjects will be requiring CPAP treatment.

The abilities of the 4-Variable screening tool, STOP, STOP-BANG, and ESS questionnaires in identifying subjects at risk for SDB were previously evaluated[37]: for predicting OSA-15, the STOP-BANG had the highest sensitivity (87.0%) with an AUC of 0.64, whereas the 4-Variable screening tool had the highest specificity (93.2%) and accuracy (79.4%). Moreover, predictive parameters for OSA-30 showed that the STOP-BANG had the highest sensitivity (70.4%), whereas the 4-Variable screening tool had the highest specificity (93.2%) and accuracy (86.7%) with an AUC of 0.67. Similar findings are shown in a study[38] comparing five different questionnaires (STOP, STOP-BANG, Berlin, ESS, and 4-Variable screening tool): the STOP-BANG had the highest sensitivity (97.6%) and the largest AUC (0.73), but the lowest specificity (12.7%) for OSA-15. Conversely, the 4-Variable screening tool had the highest specificity (74.4%) followed by ESS (67.0%).

According to NoSAS, although this instrument may possibly present good performances when other AHI thresholds are explored, it was implemented using an AHI cutoff of 20 events/h.[16] Conversely, our model was tested at three different AHI thresholds that are widely recognized and accepted when assessing the severity of OSA. The NoSAS score[16] has five variables (NC, BMI, age, sex, and snoring), whereas our tool requires only two objective parameters, a feature that we think can translate into greater practical applicability and ease of implementation. Moreover, NoSAS approach uses the presence of snoring as an integral parameter, thus requiring information based on a bed partner, thereby potentially resulting in information bias.

A recent study[39] was developed to validate the NoSAS score in a multiethnic Asian cohort and compare its performance

with STOP-BANG and Berlin questionnaires: both questionnaires performed in a manner similar to that of the NoSAS score, with AUCs of all of them clustered around 0.682–0.748. Our study also did not verify performance differences between STOP-BANG and NoSAS, therefore different from the previous study of derivation and validation of NoSAS, in which it performed better than the STOP-BANG ($P < .0001$) and Berlin questionnaires ($P < .0001$) in both cohorts (HypnoLaus and EPISONO).[16]

Strengths and Limitations

Our study had some limitations: patient selection for generation of the tool was predicated on sleep laboratory subjects, and therefore the possibility of selection bias is plausible and its implication for the general population may be limited. In general, patients referred to a sleep laboratory are often suspected of having OSA and they reflect selected patients with a high pretest probability. Furthermore, possible specific differences in the tool properties among sexes may require future optimizing adjustments. A finding that deserves to be emphasized is that our 2-item model uses a parameter (NC) that can suffer from measurement error, which can cause information bias. In addition, it can be influenced by populations with their own anthropometric characteristics (Asian or African populations), thus requiring further validation in these settings. Similarly, measures of regional obesity were not included, and might provide important additional information to the referral decision-making algorithm.

Conversely, we should also point out several features of the current study that should strengthen the ability to implement the proposed instrument. First, the tool is a mnemonic and concise model with only two numerical and readily measured objective variables along with absence of any subjectively reported items, thereby allowing for easier acquisition and calculation while reducing potential biases. Second, it was developed and then validated in two large and independent cohorts, with all individuals enrolled undergoing full PSG and with the same diagnostic criteria,[19] aiming to explore the robustness of our model. Third, the developed scoring system presented adequate performance (overall performance, discrimination, and calibration) in the derivation cohort as well as in a validation cohort.

CONCLUSIONS

In conclusion, our tool shows favorable promise for screening of OSA at any level of severity, based on commonly and widely used AHI thresholds. This tool should enable allocation of patients to different severity types and corresponding priorities, and thus enable improved patient prioritization and resource allocation. In addition, there was no superiority of one model over the other, which highlights a very great practical applicability of the No-Apnea because it contains only two objective variables easily obtained during the evaluation of a patient with suspected OSA. Therefore, it can also be used in individuals who sleep alone, in whom subjective information about sleep is not necessarily available. As with any population study, future prospective exploration for other world regions and different clinical settings will be critical for widespread implementation of such simple tool.

ABBREVIATIONS

AHI, apnea-hypopnea index
AUC, area under the curve
BMI, body mass index
CI, confidence interval
CPAP, continuous positive airway pressure
ESS, Epworth Sleepiness Scale
IQR, interquartile range
NC, neck circumference
NPV, negative predictive value
OR, odds ratio
OSA, obstructive sleep apnea
PPV, positive predictive value
PSG, polysomnography
ROC, receiver operating characteristic
SDB, sleep-disordered breathing
SpO_2, oxygen saturation
VIF, variance inflation factor

REFERENCES

1. Shah N, Roux F. The relationship of obesity and obstructive sleep apnea. *Clin Chest Med.* 2009;30(3):455–465.
2. Marshall NS, Delling L, Grunstein RR, et al. Self-reported sleep-apnoea and mortality from the Swedish Obese Subjects study. *Eur Respir J.* 2011;38(6):1349–1354.
3. Tufik S, Santos-Silva R, Taddei JA, Bittencourt LR. Obstructive sleep apnea syndrome in the Sao Paulo Epidemiologic Sleep Study. *Sleep Med.* 2010;11(5):441–446.
4. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol.* 2013;177(9):1006–1014.
5. 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.
6. Gabbay IE, Lavie P. Age- and gender-related characteristics of obstructive sleep apnea. *Sleep Breath.* 2012;16(2):453–460.
7. Duarte RL, Magalhães-da-Silveira FJ. Factors predictive of obstructive sleep apnea in patients undergoing pre-operative evaluation for bariatric surgery and referred to a sleep laboratory for polysomnography. *J Bras Pneumol.* 2015;41(5):440–448.
8. Dancey DR, Hanly PJ, Soong C, Lee B, Shepard J Jr, Hoffstein V. Gender differences in sleep apnea: the role of neck circumference. *Chest.* 2003;123(5):1544–1550.
9. Simpson L, Mukherjee S, Cooper MN, et al. Sex differences in the association of regional fat distribution with the severity of obstructive sleep apnea. *Sleep.* 2010;33(4):467–474.
10. Abrishami A, Khajehdehi A, Chung F. A systematic review of screening questionnaires for obstructive sleep apnea. *Can J Anesth.* 2010;57(5):423–438.
11. Flemons WW, Whitelaw WA, Brant R, Remmers JE. Likelihood ratios for a sleep apnea clinical prediction rule. *Am J Respir Crit Care Med.* 1994;150(5 Pt 1):1279–1285.
12. 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(7):485–491.

13. Chung F, Yegneswaran B, Liao P, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology.* 2008;108(5):812–821.
14. Takegami M, Hayashino Y, Chin K, et al. Simple four-variable screening tool for identification of patients with sleep-disordered breathing. *Sleep.* 2009;32(7):939–948.
15. Chung F, Yegneswaran B, Liao P, et al. Validation of the Berlin questionnaire and American Society of Anesthesiologists checklist as screening tools for obstructive sleep apnea in surgical patients. *Anesthesiology.* 2008;108(5):822–830.
16. 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.
17. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep.* 1991;14(6):540–545.
18. Preis SR, Massaro JM, Hoffmann U, et al. Neck circumference as a novel measure of cardiometabolic risk: the Framingham Heart Study. *J Clin Endocrinol Metab.* 2010;95(8):3701–3710.
19. 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. *J Clin Sleep Med.* 2012;8(5):597–619.
20. Yoo W, Mayberry R, Bae S, Singh K, Peter He Q, Lillard JW Jr. A study of effects of multicollinearity in the multivariable analysis. *Int J App Sci Technol.* 2014;4(5):9–19.
21. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology.* 2010;21(1):128–138.
22. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983;148(3):839–843.
23. Basoglu OK, Tasbakan MS. Gender differences in clinical and polysomnographic features of obstructive sleep apnea: a clinical study of 2827 patients. *Sleep Breath.* 2018;22(1):241–249.
24. Valipour A, Lothaller H, Rauscher H, Zwick H, Burghuber OC, Lavie P. Gender-related differences in symptoms of patients with suspected breathing disorders in sleep: a clinical population study using the sleep disorders questionnaire. *Sleep.* 2007;30(3):312–319.
25. Shah N, Hanna DB, Teng Y, et al. Sex-specific prediction models for sleep apnea from the Hispanic Community Health Study/Study of Latinos. *Chest.* 2016;149(6):1409–1418.
26. Cowan DC, Allardice G, MacFarlane D, et al. Predicting sleep disordered breathing in outpatients with suspected OSA. *BMJ Open.* 2014;4:e004519.
27. Chung F, Subramanyam R, Liao P, Sasaki E, Shapiro C, Sun Y. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth.* 2012;108(5):768–775.
28. Chung F, Yang Y, Liao P. Predictive performance of the STOP-Bang score for identifying obstructive sleep apnea in obese patients. *Obes Surg.* 2013;23(12):2050–2057.
29. Farney RJ, Walker BS, Farney RM, Snow GL, Walker JM. The STOP-Bang equivalent model and prediction of severity of obstructive sleep apnea: Relation to polysomnographic measurements of the apnea/hypopnea index. *J Clin Sleep Med.* 2011;7(5):459–465.
30. Young T, Evans L, Finn L, Palta M. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep.* 1997;20(9):705–706.
31. Stierer TL, Wright C, George A, Thompson RE, Wu CL, Collop N. Risk assessment of obstructive sleep apnea in a population of patients undergoing ambulatory surgery. *J Clin Sleep Med.* 2010;6(5):467–472.
32. Senthilvel E, Auckley D, Dasarathy J. Evaluation of sleep disorders in the primary care setting: history taking compared to questionnaires. *J Clin Sleep Med.* 2011;7(1):41–48.
33. Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. *Anesthesiology.* 2009;110(4):928–939.
34. Ahmadi N, Chung SA, Gibbs A, Shapiro CM. The Berlin questionnaire for sleep apnea in a sleep clinic population: relationship to polysomnographic measurement of respiratory disturbance. *Sleep Breath.* 2008;12(1):39–45.
35. Nagappa M, Liao P, Wong J, et al. Validation of the STOP-Bang questionnaire as a screening tool for obstructive sleep apnea among different populations: a systematic review and meta-analysis. *PLoS One.* 2015;10:e0143697.
36. Kunisaki KM, Brown KE, Fabbrini AE, Wetherbee EE, Rector TS. STOP-BANG questionnaire performance in a veterans affairs unattended sleep study program. *Ann Am Thorac Soc.* 2014;11(2):192–197.
37. 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(5):467–472.
38. Pataka A, Daskalopoulou E, Kalamaras G, Fekete Passa K, Argyropoulou P. Evaluation of five different questionnaires for assessing sleep apnea syndrome in a sleep clinic. *Sleep Med.* 2014;15(7):776–781.
39. Tan A, Hong Y, Tan LW, van Dam RM, Cheung YY, Lee CH. Validation of NoSAS score for screening of sleep-disordered breathing in a multiethnic Asian population. *Sleep Breath.* 2017;21(4):1033–1038.

ACKNOWLEDGMENTS

We thank José Roberto Lapa e Silva, MD, PhD for critical review of the manuscript.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication October 19, 2017
Submitted in final revised form February 14, 2018
Accepted for publication February 23, 2018
Address correspondence to: Ricardo L. M. Duarte. Sleep - Laboratório de Estudo dos Distúrbios do Sono, Centro Médico BarraShopping, Avenida das Américas 4666, sala 309, Barra da Tijuca, 22649-900, Rio de Janeiro, Brazil; Tel: 55 21 2430-9222; Fax: 55 21 2430-9220; Email: rlmduarte@gmail.com

DISCLOSURE STATEMENT

All authors have seen and approved the manuscript. The authors report no conflicts of interest.