Nature and Science of Sleep

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Obstructive Sleep Apnea Screening with a 4-Item Instrument, Named GOAL Questionnaire: Development, Validation and Comparative Study with No-Apnea, STOP-Bang, and NoSAS

This article was published in the following Dove Press journal:
Nature and Science of Sleep

Ricardo LM Duarte[1,2]
Flavio J Magalhães-da-Silveira[1]
Tiago S Oliveira-e-Sá[3,4]
Joana A Silva[5]
Fernanda CQ Mello[2]
David Gozal[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]Hospital de Santa Marta, Centro Hospitalar Lisboa Central, Lisbon, Portugal; [4]NOVA Medical School, Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisbon, Portugal; [5]Clínica São Vicente, Rede D'Or, Rio de Janeiro, Brazil; [6]Department of Child Health, University of Missouri School of Medicine, Columbia, MO, USA

Background: Obstructive sleep apnea (OSA) is a very prevalent disorder. Here, we aimed to develop and validate a practical questionnaire with yes-or-no answers, and to compare its performance with other well-validated instruments: No-Apnea, STOP-Bang, and NoSAS.

Methods: A cross-sectional study containing consecutively selected sleep-lab subjects underwent full polysomnography. A 4-item model, named GOAL questionnaire (gender, obesity, age, and loud snoring), was developed and subsequently validated, with item-scoring of 0–4 points (≥2 points indicating high risk for OSA). Discrimination was assessed by area under the curve (AUC), while predictive parameters were calculated using contingency tables. OSA severity was classified based on conventionally accepted apnea/hypopnea index thresholds: ≥5.0/h ($OSA_{\geq 5}$), ≥15.0/h ($OSA_{\geq 15}$), and ≥30.0/h ($OSA_{\geq 30}$).

Results: Overall, 7377 adults were grouped into two large and independent cohorts: derivation (n = 3771) and validation (n = 3606). In the derivation cohort, screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$ revealed that GOAL questionnaire achieved sensitivity ranging from 83.3% to 94.0% and specificity ranging from 62.4% to 38.5%. In the validation cohort, screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$, corroborated validation steps with sensitivity ranging from 83.7% to 94.2% and specificity from 63.4% to 37.7%. In both cohorts, discriminatory ability of GOAL questionnaire for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$ was similar to No-Apnea, STOP-Bang or NoSAS.

Conclusion: All four instruments had similar performance, leading to a possible greater practical implementation of the GOAL questionnaire, a simple instrument with only four parameters easily obtained during clinical evaluation.

Keywords: obstructive sleep apnea, polysomnography, screening, questionnaire, diagnosis

Introduction

Obstructive sleep apnea (OSA) is characterized by frequent partial or complete upper airway collapse during sleep, resulting in intermittent hypoxemia and sleep fragmentation.[1] There is growing evidence that OSA plays a key role in the pathogenesis of cardiovascular and metabolic diseases, as well as a clear association with increased overall mortality and medical costs.[2,3] OSA is very prevalent disease,[4–6] being that most recent estimates have suggested that nearly a billion of people may be affected, with about 425 million individuals, aged 30 to 69, suffering from moderate to severe OSA, worldwide.[7]

Correspondence: David Gozal
Department of Child Health, University of Missouri School of Medicine, Columbia, MO, USA
Email drdavidgozal@gmail.com

submit your manuscript | www.dovepress.com
DovePress
http://doi.org/10.2147/NSS.S238255

The gold standard for the diagnosis of OSA is an attended in-lab overnight polysomnography (PSG) test.[1] However, this costly method is not readily available to a large number of patients with suspected OSA, especially in regions with limited economic resources.[1,8] Due to the high frequency of OSA, there is a significant cost in evaluating all subjects with suspected OSA, leading to long waiting times for testing.[9] In this context, portable diagnostic methods have been extensively implemented as an alternative for the diagnosis of OSA in those adults exhibiting a high pretest probability and absence of significant comorbidity.[1]

The various screening instruments that have merged over the years have been generally predicated on the clinical, demographic and anthropometric parameters that constitute the major risk factors for OSA.[8] The STOP-Bang questionnaire is an 8-item acronym that can be easily completed.[10] It was developed and later validated in surgical subjects having a high prevalence of $OSA_{\geq 5}$: 73% (derivation cohort) and 69% (validation cohort).[10] Subsequently, it was widely validated in several different settings: sleep clinic patients, surgical patients, and general population.[11–16] Conversely, NoSAS score was derived and validated in population-based cohorts (HypnoLaus and EPISONO, respectively).[17] Afterwards, this score was also validated in different settings, reporting adequate performance as screening model for OSA: in a multiethnic Asian cohort, in a hospital-based sample, in depressive subjects, and in a sleep clinic.[18–22]

The No-Apnea[23] tool is a newly developed and validated screening instrument that includes only two objective parameters: neck circumference (NC) and age, with a total score ranging from 0 to 9 (cutoff point ≥3 classifies patients at high risk for OSA). In derivation and validation cohorts, No-Apnea showed adequate performance in OSA screening, with a discriminatory ability that was similar and indistinguishable from the performance of either STOP-Bang or NoSAS.[23] Although No-Apnea has been subsequently validated in selected clinical populations,[24–26] and displayed adequate predictive performance, some points deserve to be highlighted: i) absence of subjective variables may result in more difficult applicability, especially in regions with a lower prevalence of OSA, such as in primary care; ii) presence of age ≥55 years, already sufficient to classify individuals at high risk for the diagnosis of OSA, possibly limits its use in older populations; and iii) because it contains different scores for each parameter, it may be difficult to adopt and implement, particularly in general clinical practice settings.

Based on such concerns, we aimed to develop a practical and concise questionnaire that would include both objective and subjective variables, with yes-or-no dichotomous answers. To this effect, we designed the present study into two parts: i) derivation and validation of a proposed tool for screening OSA in adults, and ii) comparison with three other widely validated OSA screening instruments: No-Apnea score, STOP-Bang questionnaire, and NoSAS score.

Methods

Study Design and Patient Selection

This was a cross-sectional study, comprising the period from January 2017 to June 2019, including adults which were consecutively referred for PSG evaluation due to suspected sleep disordered breathing by their attending physicians. Then, all subjects were grouped into two separate cohorts: derivation (from January 2017 to February 2018) and validation (from May 2018 to June 2019). Inclusion criteria consisted of individuals of both genders, aged ≥18 years and with suspected of OSA, while exclusion criteria were: previously diagnosed OSA, use of home sleep study for diagnosis, incomplete clinical data and technically inadequate PSG. In case the same individual underwent more than one PSG, only the test with the longest total sleep time was included. All sleep tests were conducted in a Brazilian single-center located in the city of Rio de Janeiro: SLEEP – Laboratório de Estudo dos Distúrbios do Sono.

On the evening of PSG, clinical, demographic and anthropometric data were systematically obtained from all participants by trained sleep laboratory technicians: gender, age, body mass index (BMI), NC, self-reported comorbidities (hypertension and diabetes mellitus), OSA-related symptoms (snoring, observed apnea, gasping/choking, and tiredness), and Epworth Sleepiness Scale (ESS).[27] The ESS is an 8-item questionnaire (each item is scored from 0 to 3 points, with a final score from 0 to 24 points), which subjectively measures excessive daytime sleepiness by a score ≥11 points.[27] BMI was calculated dividing the weight in kilograms by the square of the height in meters (kg/m^2), while NC (cm) was measured using a tape measure, with its upper edge placed below the laryngeal prominence and applied perpendicularly along the neck axis.

Ethical Considerations

The study protocol (No. 1.764.165) was approved by the Research Ethics Committee of the Federal University of Rio de Janeiro (UFRJ) and adhered to standard previously set by the Declaration of Helsinki. All participants

provided written informed consent, being that anonymity of all recruited individuals was guaranteed throughout the study process.

Screening Instruments

The No-Apnea (final score from 0 to 9, high risk with 3 or more points) contains two objective parameters: NC is scored in three values: 1 (37.0–39.9 cm), 3 (40.0–42.9 cm), and 6 (-≥43.0 cm), while age is scored as follows: 1 (35–44 years), 2 (45–54 years), and 3 (≥55 years).[23] The STOP-Bang (final score from 0 to 8, high risk with 3 or more points) consists of 8 yes-or-no questions (1 point for each affirmative answer): loud snoring, tiredness, observed apnea, hypertension, BMI > 35 kg/m^2, age >50 years, NC > 40 cm, and male gender.[10] The NoSAS allocates 4 points for having a NC >40 cm, 3 points for having a BMI of 25–29 kg/m^2 or 5 points for having a BMI ≥ 30 kg/m^2, 2 points for habitual snoring, 4 points for age >55 years, and 2 points for men; this instrument is considered positive with a score ≥8 points (from 0 to 17 points).[17]

Sleep Studies

All subjects underwent an attended, in-lab PSG (EMBLA® S7000, Embla Systems, Inc., Broomfield, Colorado, United States) consisting of continuous monitoring of electroencephalogram, electrooculogram, electromyogram (chin and legs), electrocardiogram, airflow, thoracic and abdominal impedance belts, oxygen saturation, snoring microphone and sensors for body position. Polysomnographic records were manually interpreted by two board-certified sleep physicians, according to a guideline previously published in 2012 by the American Academy of Sleep Medicine (AASM),[28] which were blinded to the values of all screening instruments collected prior to PSG. Apneas were classified from a drop ≥90% of baseline airflow lasting at least 10 s, while hypopneas were classified from a ≥30% pre-event drop over ≥10 s associated with desaturation of oxygen ≥3% or an arousal.[28] The AHI was calculated as the number of apnea plus hypopnea/total sleep time (in hours). Polysomnographic diagnosis of OSA was based on apnea/hypopnea index (AHI) ≥5.0/h and its severity was classified as follows: ≥5.0/h as any OSA (OSA$_{\geq 5}$), ≥15.0/h as moderate/severe OSA (OSA$_{\geq 15}$), and ≥30.0/h as severe OSA (OSA$_{\geq 30}$).

Statistical Analysis

Data analysis was carried out using SPSS for Windows (version 21.0; SPSS; Chicago, IL, United States). Results are presented as mean ± standard deviation for continuous variables and as frequency (n) with percentage (%) for categorical variables. The groups were compared using the chi-square test for categorical variables, while Student's *t*-test and univariate analysis of variance (ANOVA) were used for numerical variables. The modeling of the proposed instrument was initially based on 10 variables often associated with OSA diagnosis: male gender, age ≥50 years, NC ≥40 cm, BMI ≥30 kg/m^2, loud snoring, observed apnea, hypertension, diabetes mellitus, tiredness, and gasping/choking. The chosen cutoff points for age and BMI were based on the highest odds ratio (OR) obtained for OSA$_{\geq 15}$ diagnosis: age ≥50 years (OR: 1.911; 95% confidence interval [CI]: 1.668–2.189) versus age ≥55 years (OR: 1.817; 95% CI: 1.566–2.108) and BMI ≥ 30 kg/m^2 (OR: 2.563; 95% CI: 2.-243–2.928) versus BMI ≥ 35 kg/m^2 (OR: 2.088; 95% CI: 1.826–2.388). Similarly, we used loud snoring and not the presence of snoring by obtaining a higher OR for OSA$_{\geq 15}$ diagnosis: OR: 5.454 (95% CI: 3.972–7.490) versus 3.801 (95% CI: 3.300–4.377). The NC ≥ 40 cm has been chosen as this is the threshold used in the STOP-Bang and NoSAS instruments.[10,17] Firstly, these 10 chosen variables were evaluated by univariate analyses, where the outcome was an AHI ≥ 15.0/h, a conventionally accepted cutoff point for classifying clinically relevant OSA. Sequentially, clinically relevant variables ($p < 0.10$ in the univariate analysis) were included in a multivariate logistic regression model with AHI ≥ 15.0/h being the dependent variable. Performance of the newly developed instrument was calculated as follows: i) discrimination by the Receiver Operating Characteristic (ROC) curves and area under the curve (AUC); ii) calibration assessed by the Hosmer–Lemeshow test ($p < 0.05$ was considered as poor calibration); iii) 2x2 contingency tables (sensitivity, specificity, positive predictive value [PPV], and negative predictive value [NPV]); and iv) correlation through the Spearman coefficient (r_s). As Hosmer–Lemeshow test is very sensitive to sample size, we chose a smaller subset of randomly selected patients (n = 1000) to evaluate calibration. An AUC > 0.7 was considered as clinically significant, being that AUCs were compared using a previously described algorithm.[29] All predictive parameters were calculated at three AHI thresholds (5.0/h, 15.0/h, and 30.0/h) and reported with their respective 95% CIs. A two-tailed p-value <0.05 was considered as statistically significant.

Results

The flowchart of the study is showed in Figure 1, being that 7377 subjects were consecutively allocated into two independent cohorts: one for derivation (n = 3771) and one for validation (n = 3606). According to Table 1, no clinical or polysomnographic parameter showed a statistically significant difference between derivation and validation cohorts, except for total sleep time (p = 0.047) and arousal index (p = 0.031).

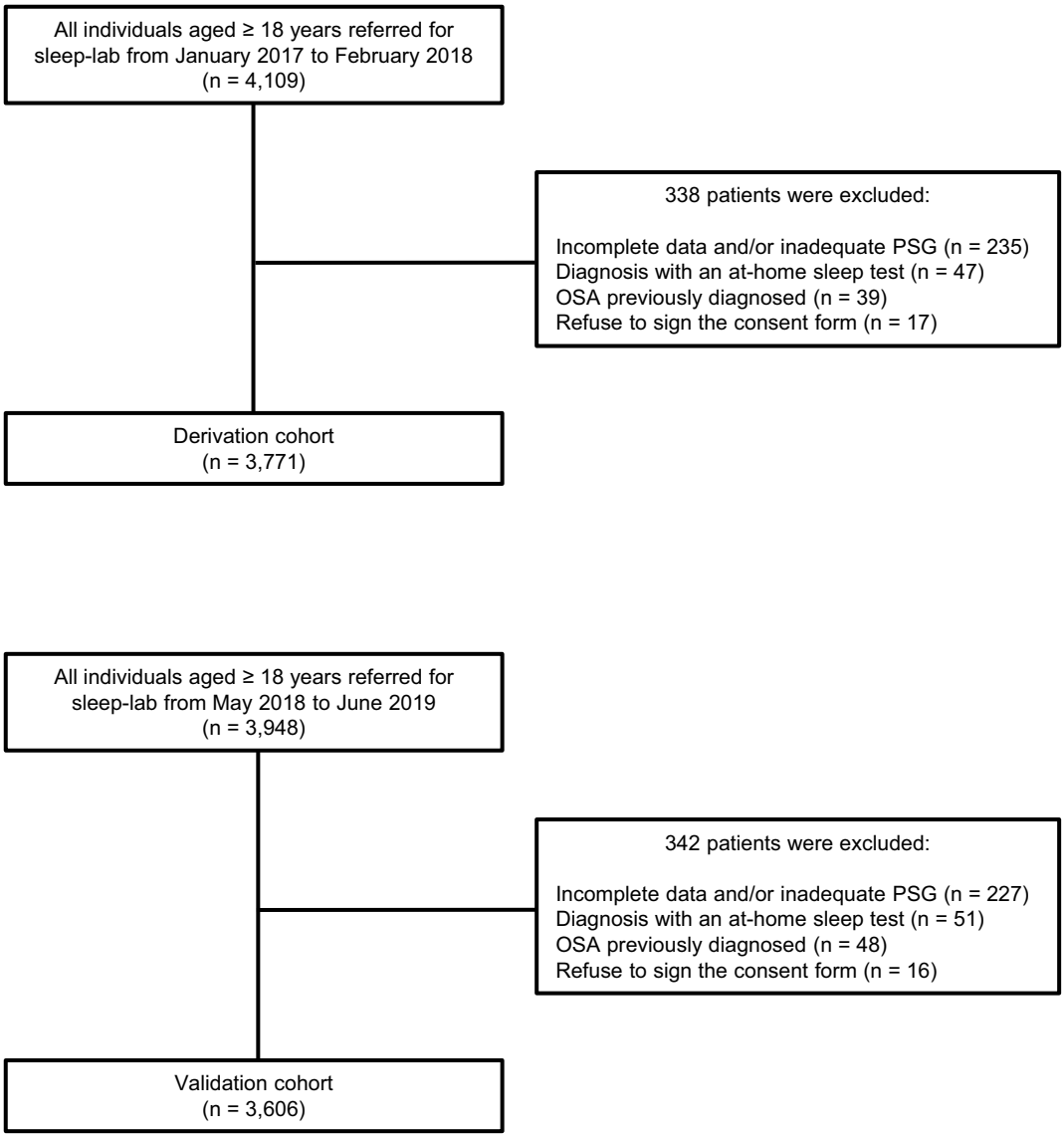

Figure 1 The flowchart of the patients.
Abbreviations: OSA, obstructive sleep apnea; PSG, polysomnography.

Prevalence of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$ showed no statistically significant differences between derivation and validation cohorts: 79.1% versus 78.8% (p = 0.753), 57.7% versus 57.0% (p = 0.604), and 37.8% versus 35.7% (p = 0.070); respectively.

Modeling

Table 2 shows the seven independent predictive parameters for screening for $OSA_{\geq 15}$. Diabetes mellitus and gasping/choking, although initially tested as possible OSA predictors, were subsequently excluded because they are not independent variables for $OSA_{\geq 15}$ diagnosis. As can be seen in Table 2, these seven parameters were ranked according to adjusted OR: the main parameter was male gender, followed by age ≥50 years, loud snoring, BMI ≥ 30 kg/m^2, NC ≥ 40 cm, observed apnea, and hypertension.

From these data, several possible models were tested using the main clinical parameters previously found. Figure 2 summarizes the discrimination of possible models containing from 3 to 7 parameters, which exemplifies why a 4-item model was ultimately chosen, since discriminatory power did not significantly increase with the addition of a 5th (NC ≥ 40 cm), 6th (observed apnea) or 7th variable (hypertension). This 4-item instrument, because it contains only yes-or-no dichotomous answers, reaches a final score of 0–4 points (Table 3), being later named by the acronym GOAL (gender, obesity, age, and loud snoring) questionnaire.

Table 1 Patient Characteristics

Parameters	Derivation Cohort (n = 3771)	Validation Cohort (n = 3606)	p
Clinical data			
Male gender (%)	1983 (52.6)	1961 (54.4)	0.123
Age (years)	45.9 ± 14.6	45.7 ± 14.6	0.593
BMI (kg/m^2)	33.1 ± 7.9	32.9 ± 7.7	0.437
NC (cm)	40.5 ± 5.0	40.5 ± 4.8	0.998
ESS (points)	9.8 ± 5.0	9.8 ± 5.1	0.766
Hypertension (%)	1531 (40.6)	1404 (38.9)	0.147
Diabetes mellitus (%)	470 (12.5)	436 (12.1)	0.645
Loud snoring (%)	2465 (65.4)	2291 (63.5)	0.103
Observed apnea (%)	1893 (50.2)	1847 (51.2)	0.389
Gasping/choking (%)	1503 (39.9)	1488 (41.3)	0.226
Tiredness (%)	2711 (71.9)	2607 (72.3)	0.716
Polysomnographic data			
Total sleep time (min)	338.5 ± 71.7	342.1 ± 68.2	0.047
NREM sleep stage (%)	83.6 ± 7.8	83.4 ± 7.8	0.296
REM sleep stage (%)	15.9 ± 7.8	16.2 ± 7.9	0.179
Arousal index (n/h)	31.1 ± 26.0	29.8 ± 24.8	0.031
AHI (n/h)	28.9 ± 28.1	27.7 ± 26.9	0.058
Mean SpO_2 (%)	93.5 ± 3.2	93.4 ± 3.5	0.528
Nadir SpO_2 (%)	81.8 ± 9.3	81.9 ± 9.4	0.680
AHI (n/h)			
≥5.0/h (%)	2984 (79.1)	2842 (78.8)	0.753
≥15.0/h (%)	2174 (57.7)	2057 (57.0)	0.604
≥30.0/h (%)	1424 (37.8)	1288 (35.7)	0.070

Note: Numeric and categorical variables were reported as mean ± SD and n (%), respectively.

Abbreviations: BMI, body mass index; NC, neck circumference; ESS, Epworth Sleepiness Scale; NREM, non-rapid eye movement; REM, rapid eye movement; AHI, apnea/hypopnea index; SpO_2, oxygen saturation.

Predictive Parameters

Mean AHI values increased linearly in parallel with increasing GOAL questionnaire scores (from 0 to 4 points): i) derivation cohort: from 3.6 ± 5.7/h to 53.3 ± 26.1/h (p-value for trend <0.001) and ii) validation cohort: from 3.6 ± 5.8/h to 50.8 ± 27.1/h (p-value for trend <0.001); Figure 3. In both datasets, GOAL was positively correlated with AHI: $r_s = 0.559$ with $p < 0.001$ (derivation cohort) and $r_s = 0.554$ with $p < 0.001$ (validation cohort). In the derivation cohort, GOAL questionnaire showed adequate calibration for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$ assessed by Hosmer–Lemeshow test: 11.563 (p = 0.172), 5.135 (p = 0.743), and 5.355 (p = 0.719); respectively. Similarly, in the validation cohort, GOAL displayed adequate calibration for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$ assessed by Hosmer–Lemeshow test: 9.855 (p = 0.197), 3.921 (p = 0.864), and 5.799 (p = 0.670); respectively.

Table 2 Univariate and Multivariate Analyses of Clinical Parameters According to AHI ≥ 15.0/h (Derivation Cohort; n = 3771)

Parameters	Univariate Analysis		Multivariate Analysis	
	Unadjusted OR (95% CI)	p	Adjusted OR (95% CI)	p
Male gender	3.394 (2.964–3.885)	< 0.001	2.831 (2.356–3.402)	< 0.001
Age ≥ 50 years	1.911 (1.668–2.189)	< 0.001	2.572 (2.159–3.064)	< 0.001
Loud snoring	5.454 (3.972–7.490)	< 0.001	2.402 (2.037–2.832)	< 0.001
BMI ≥ 30 kg/m^2	2.563 (2.243–2.928)	< 0.001	2.315 (1.932–2.774)	< 0.001
NC ≥ 40 cm	4.323 (3.759–4.972)	< 0.001	1.937 (1.602–2.343)	< 0.001
Observed apnea	2.894 (2.531–3.309)	< 0.001	1.670 (1.412–1.975)	< 0.001
Hypertension	2.147 (1.874–2.459)	< 0.001	1.254 (1.060–1.484)	0.008
Diabetes mellitus	1.841 (1.494–2.269)	< 0.001	1.140 (0.892–1.457)	0.295
Gasping/choking	1.342 (1.175–1.533)	< 0.001	1.022 (0.867–1.204)	0.799
Tiredness	0.913 (0.790–1.054)	0.226	–	–

Abbreviations: AHI, apnea/hypopnea index; BMI, body mass index; NC, neck circumference; OR, odds ratio; CI, confidence interval.

Predictive performance of the GOAL questionnaire is shown in Table 4 (derivation cohort). We used a cutoff point ≥2 to classify patients at high risk for OSA. This strategy focused on privileging sensitivity over specificity, aiming at reducing the false-negative rate. The frequency of subjects classified as high risk for OSA through the four screening instruments is shown in Figure 4.

Table 5 compares GOAL questionnaire with three other previously validated tools: No-Apnea, STOP-Bang, and NoSAS. In the derivation cohort, for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$, GOAL questionnaire presented sensitivities ranging from 83.3% to 94.0% and specificities ranging from 62.4% to 38.5%. In the validation cohort, for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$, GOAL questionnaire showed sensitivity values ranging from 83.7% to 94.2% and specificity values ranging from 63.4% to 37.7%. In both cohorts, for diagnosis of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$, STOP-Bang always presented the highest sensitivity, while NoSAS always presented the highest specificity. In both datasets and at all OSA severity levels, GOAL questionnaire showed high sensitivity. As expected, for all four instruments, as AHI thresholds increased, a corresponding increase in sensitivity and a reduction in specificity occurred.

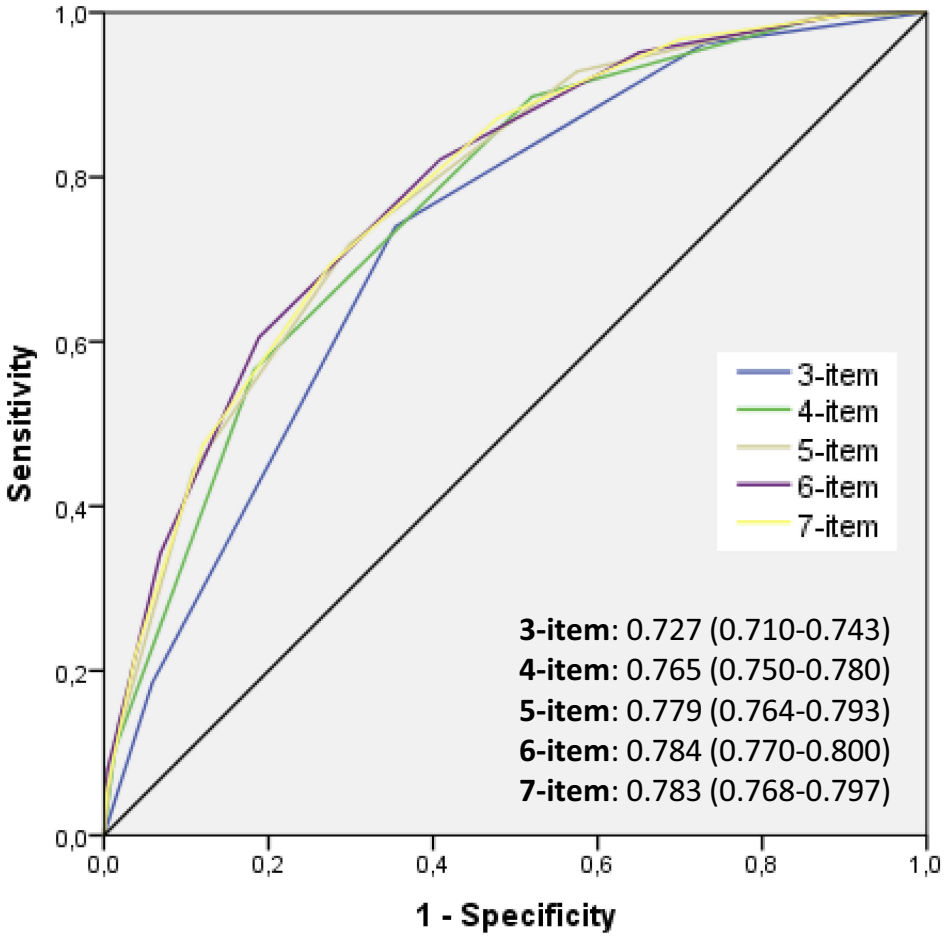

Figure 2 Graphical representation of several models sequentially constructed with the main independent variables for screening of moderate/severe obstructive sleep apnea ($OSA_{\geq 15}$). 3-item model: male gender, age ≥50 years, and loud snoring; 4-item model: 3-item model plus body mass index ≥30 kg/m^2, 5-item model: 4-item model plus neck circumference ≥40 cm, 6-item model: 5-item model plus observed apnea, and 7-item model: 6-item model plus hypertension. The 3-item model reported discrimination statistically lower than that obtained by 4-item model ($p < 0.001$). The 4-item model had a similar discrimination when compared to the 5-item, 6-item or 7-item models (p-values: 0.184, 0.070, and 0.086; respectively). Estimates reported as area under the curve (95% confidence interval).

Comparing Discriminatory Ability

Figure 5 shows the discrimination obtained by the four screening instruments: in both cohorts and at all OSA severity levels, there were no statistically significant differences in the discriminatory ability of the GOAL questionnaire when compared to No-Apnea, STOP-Bang or NoSAS in screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$.

Table 3 The GOAL Questionnaire

Parameters	Points
G - Male gender	No = 0 Yes = 1
O - Obesity: body mass index ≥ 30 kg/m^2	No = 0 Yes = 1
A - Age ≥ 50 years	No = 0 Yes = 1
L - Loud snoring	No = 0 Yes = 1

Note: The points for each variable are added, totaling a final score of 0–4 points.

Discussion

Current findings show that in a sleep-lab setting, a simple and practical 4-item model, which we have denominated GOAL questionnaire (male gender, obesity with BMI ≥30 kg/m^2, age ≥50 years, and loud snoring), achieves adequate and reproducible predictive performance for OSA screening at all severity levels. Furthermore, its discriminatory ability was similar from that of three previously validated and widely used screening instruments for OSA: No-Apnea score, STOP-Bang questionnaire, and NoSAS score. Moreover, obtaining score ≥2 versus <2 points to rank, respectively, at high or low risk allowed an increase in specificity values compared to No-Apnea, while concurrently preserving a relatively high sensitivity, especially at most severe levels.

Sensitivity and specificity of a screening model are usually inversely related and high sensitivity often comes

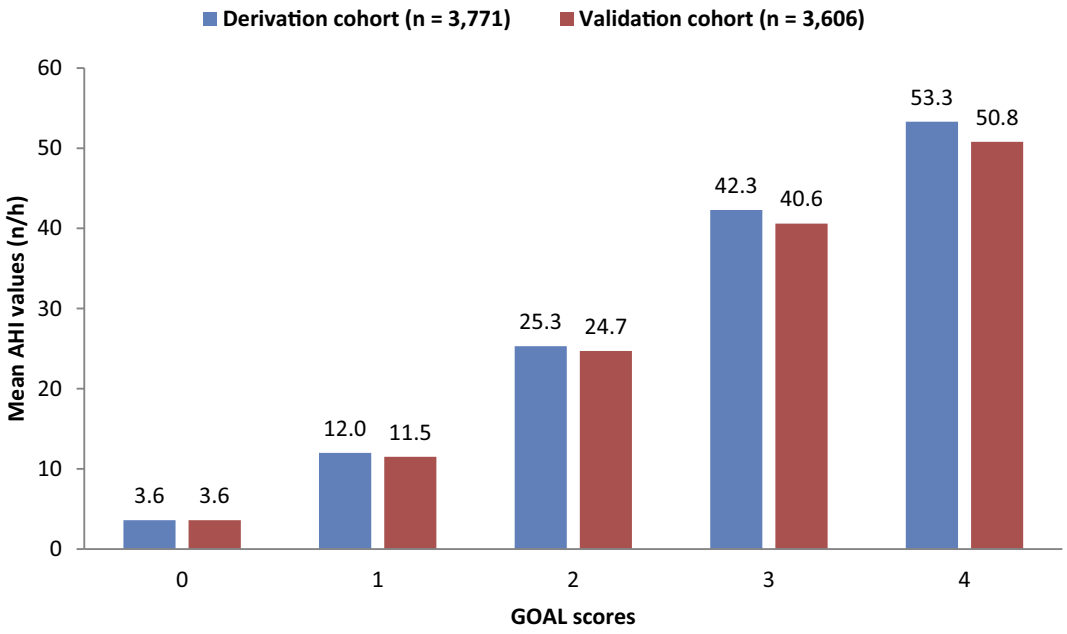

Figure 3 Mean apnea/hypopnea index (AHI) values obtained by polysomnography according to GOAL questionnaire scores (from 0 to 4 points).

at the expense of specificity.[8,30] In the context of OSA, which is a very prevalent disease and often associated with significant morbidity, it may be more important that the screening method has high sensitivity, with consequent reduction in the false-negative rate.[8,30] As expected, all four screening instruments (GOAL, No-Apnea, STOP-Bang, and NoSAS) showed increased sensitivity in parallel with increasing AHI thresholds (from 5.0/h to 30.0/h). On the other hand, specificity decreased in the most severe forms of OSA, which can result in a considerable number of false positives.

Although our newly developed questionnaire contains twice as many parameters as No-Apnea, it may be easier to respond, score and interpret, since its construct is predicated

Table 4 Predictive Parameters of the GOAL Questionnaire (Derivation Cohort; n = 3771)

	GOAL Questionnaire Scores			
	≥ 1 versus < 1	**≥ 2 versus < 2**	**≥ 3 versus < 3**	**4 versus < 4**
AHI ≥ 5.0/h				
Sensitivity	98.4 (98.0–98.8)	83.3 (82.6–84.1)	48.0 (47.4–48.5)	8.20 (8.00–8.30)
Specificity	20.5 (18.8–21.8)	62.4 (59.5–65.2)	88.9 (86.6–90.9)	99.5 (98.6–99.8)
PPV	82.4 (82.1–82.7)	89.4 (88.5–90.2)	94.3 (93.1–95.3)	98.4 (95.7–99.5)
NPV	77.4 (71.2–82.6)	49.7 (47.4–52.0)	31.1 (30.3–31.8)	22.2 (22.0–22.3)
AHI ≥ 15.0/h				
Sensitivity	99.6 (99.2–99.8)	89.8 (88.7–90.8)	56.5 (55.3–57.7)	10.5 (10.0–10.8)
Specificity	12.5 (11.9–12.7)	48.0 (46.5–49.4)	81.8 (80.1–83.3)	98.6 (97.9–99.1)
PPV	60.8 (60.5–60.9)	70.1 (69.3–70.9)	80.9 (79.1–82.5)	91.2 (86.9–94.3)
NPV	95.7 (91.7–97.9)	77.5 (75.1–79.8)	58.0 (56.9–59.1)	44.7 (44.4–44.9)
AHI ≥ 30.0/h				
Sensitivity	99.9 (99.4–100.0)	94.0 (92.8–95.1)	65.8 (63.9–67.7)	13.6 (12.6–14.5)
Specificity	8.80 (8.50–8.89)	38.5 (37.7–39.1)	75.2 (74.0–76.3)	97.6 (97.0–98.1)
PPV	39.9 (39.7–40.0)	48.1 (47.5–48.7)	61.6 (59.8–63.4)	77.6 (72.0–82.4)
NPV	99.0 (96.2–99.8)	91.4 (89.6–93.0)	78.4 (77.1–79.6)	65.1 (64.7–65.4)

Note: Data are presented as estimates (95% confidence intervals).
Abbreviations: AHI, apnea/hypopnea index; PPV, positive predictive value; NPV, negative predictive value.

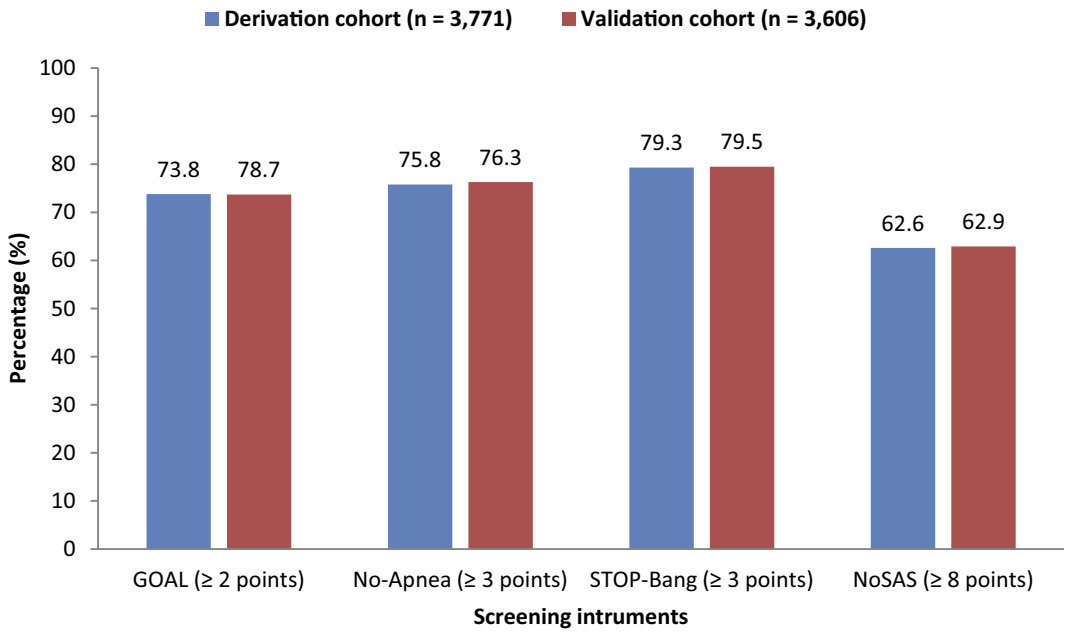

Figure 4 Percentage of individuals assessed as high risk for diagnosis of obstructive sleep apnea by four screening instruments: GOAL questionnaire, No-Apnea score, STOP-Bang questionnaire, and NoSAS score.

on yes-or-no answers leading to easy calculation of the cumulative score and simple attribution of risk. In contrast, No-Apnea contains several sub-items being scored from 0 to 9 points. Another noteworthy fact is that No-Apnea derivation and validation study[23] was designed with retrospectively recruited subjects, while the present study was conducted with two independent cohorts of prospectively enrolled subjects. In addition, logistic regression employed for No-Apnea model development[23] was performed with binary outcome (AHI ≥ 5.0/h), while in the development of the GOAL questionnaire we employed AHI ≥ 15.0/h, a threshold often implicated as a clinically relevant demarcation for OSA. Our recently proposed instrument contains only one subjective variable (loud snoring), similar to NoSAS (presence of snoring), but well below the STOP-Bang which has four subjective variables (loud snoring, observed apnea, tiredness, and hypertension). Using a limited number of subjective parameters possibly could result in greater applicability of the instrument and less possibility of bias.

Another feature of the current instrument is that by increasing its scores (from 0 to 4 points) there was a linear increase in AHI obtained by PSG. This property may further translate into a valuable tool for health professionals in defining the level of risk and relative certainty when attempting to identify those patients with a higher or lower probability of having a diagnosis of OSA. The risk escalation characteristic has also been previously described with the STOP-Bang questionnaire, whereby increasing values obtained from STOP-Bang (from 0 to 8 points) were associated with an increased likelihood of having OSA.[31]

The four clinical parameters (gender, BMI, age, snoring) employed for the elaboration of our screening instrument are well-established predictive factors for OSA. Regarding clinical symptomatology, men with OSA usually have typical symptoms such as snoring and observed apnea, while women with OSA often report symptoms that are considered atypical, such as insomnia, morning headache, and fatigue.[32,33] Based on polysomnographic findings, women have a lower prevalence of OSA than men, but with a clear increase in frequency after menopause.[32,33] Unlike No-Apnea,[23] the chosen measure of obesity was BMI over NC: this finding may translate into a better balance of its components between genders, since NC is a variable that is arguably higher in men than in women.[32–34] OSA prevalence increases with age, with some studies reporting a high prevalence of OSA in the elderly.[35] The only subjective variable that was chosen in our model was loud snoring, since snoring intensity increases as OSA becomes more severe.[36] In addition, this question from the STOP-Bang questionnaire "Do you snore loudly?" has already undergone cross-cultural adaptation to the Portuguese language spoken in Brazil.[37]

Limitations and Strengths

Our study has some limitations that should be highlighted: the study population was based on sleep-lab referral patients

Table 5 Predictive Parameters of the OSA Screening Models

	Screening Instruments			
	GOAL	**No-Apnea**	**STOP-Bang**	**NoSAS**
Derivation cohort				
AHI ≥ 5.0/h				
Sensitivity	83.3 (82.6–84.1)	84.3 (83.5–85.0)	87.2 (86.5–88.0)	71.5 (70.7–72.3)
Specificity	62.4 (59.5–65.2)	56.3 (53.4–59.2)	50.6 (47.7–53.4)	71.0 (68.1–73.8)
PPV	89.4 (88.5–90.2)	88.0 (87.2–88.8)	87.0 (86.2–87.7)	90.3 (89.4–91.3)
NPV	49.7 (47.4–52.0)	48.6 (46.0–51.0)	51.1 (48.2–53.9)	39.7 (38.0–41.2)
AHI ≥ 15.0/h				
Sensitivity	89.8 (88.7–90.8)	89.8 (88.7–90.9)	91.7 (90.7–92.6)	79.0 (77.7–80.2)
Specificity	48.0 (46.5–49.4)	43.3 (41.8–44.7)	37.4 (36.1–38.7)	59.6 (57.9–61.3)
PPV	70.1 (69.3–70.9)	68.3 (67.5–69.1)	66.6 (65.9–67.3)	72.7 (71.5–73.8)
NPV	77.5 (75.1–79.8)	75.8 (73.2–78.2)	76.8 (73.9–79.4)	67.6 (65.6–69.5)
AHI ≥ 30.0/h				
Sensitivity	94.0 (92.8–95.1)	93.2 (91.9–94.3)	95.6 (94.4–96.5)	85.3 (83.6–86.8)
Specificity	38.5 (37.7–39.1)	34.7 (33.9–35.4)	30.5 (29.8–31.1)	51.1 (50.1–52.1)
PPV	48.1 (47.5–48.7)	46.4 (45.8–47.0)	45.5 (44.9–45.9)	51.4 (50.4–52.4)
NPV	91.4 (89.6–93.0)	89.4 (87.3–91.2)	91.9 (89.8–93.6)	85.1 (83.4–86.7)
Validation cohort				
AHI ≥ 5.0/h				
Sensitivity	83.7 (82.9–84.4)	85.2 (84.4–86.0)	88.0 (87.2–88.8)	71.8 (71.0–72.6)
Specificity	63.4 (60.4–66.2)	56.8 (53.9–59.7)	52.1 (49.2–54.9)	70.2 (67.2–73.0)
PPV	89.5 (88.6–90.3)	88.0 (87.2–88.8)	87.2 (86.5–88.0)	89.9 (88.9–90.9)
NPV	51.1 (48.7–53.3)	50.8 (48.1–53.3)	53.9 (50.9–56.7)	40.1 (38.3–41.7)
AHI ≥ 15.0/h				
Sensitivity	89.2 (88.0–90.2)	90.4 (89.3–91.5)	92.5 (91.5–93.5)	79.0 (77.7–80.3)
Specificity	46.8 (45.3–48.2)	42.5 (41.0–43.9)	37.8 (36.4–39.0)	58.5 (56.7–60.2)
PPV	69.0 (68.1–69.8)	67.6 (66.8–68.4)	66.4 (65.6–67.1)	71.6 (70.5–72.8)
NPV	76.5 (74.0–78.8)	77.0 (74.3–79.5)	79.2 (76.3–81.8)	67.7 (65.7–69.7)
AHI ≥ 30.0/h				
Sensitivity	94.2 (92.8–95.3)	93.8 (92.4–95.0)	96.6 (95.5–97.5)	85.7 (83.9–87.4)
Specificity	37.7 (36.9–38.3)	33.4 (32.7–34.1)	30.0 (29.4–30.5)	49.8 (48.8–50.7)
PPV	45.6 (45.0–46.2)	43.9 (43.3–44.5)	43.4 (42.9–43.8)	48.7 (47.7–49.6)
NPV	92.1 (90.3–93.6)	90.6 (88.6–92.4)	94.0 (92.1–95.6)	86.2 (84.5–87.9)

Note: Data are presented as estimates (95% confidence intervals).
Abbreviations: AHI, apnea/hypopnea index; PPV, positive predictive value; NPV, negative predictive value.

(subjects with a high pretest probability), and therefore the possibility of selection bias is plausible. As the predictive values of a screening model are affected by disease prevalence,[30,38] future studies should be conducted to assess the performance of the GOAL questionnaire in other population groups. In addition, it was conducted in a single-center and did not preferably include individuals belonging to populations with their own distinct anthropometric characteristics (eg, Asian or African populations), thus requiring future additional external validation. Conversely, we should also point out several features of the present study that should strengthen the ability to implement the proposed instrument: it was developed and validated from two very large and representative cohorts, with prospectively recruited participants undergoing the same diagnostic test for OSA (overnight full PSG) and with the same diagnostic criteria recommended by AASM.[28]

Conclusions

All instruments (GOAL questionnaire, No-Apnea score, STOP-Bang questionnaire and NoSAS score) were found to be adequate instruments for OSA screening at any severity level. The use of these instruments can allow

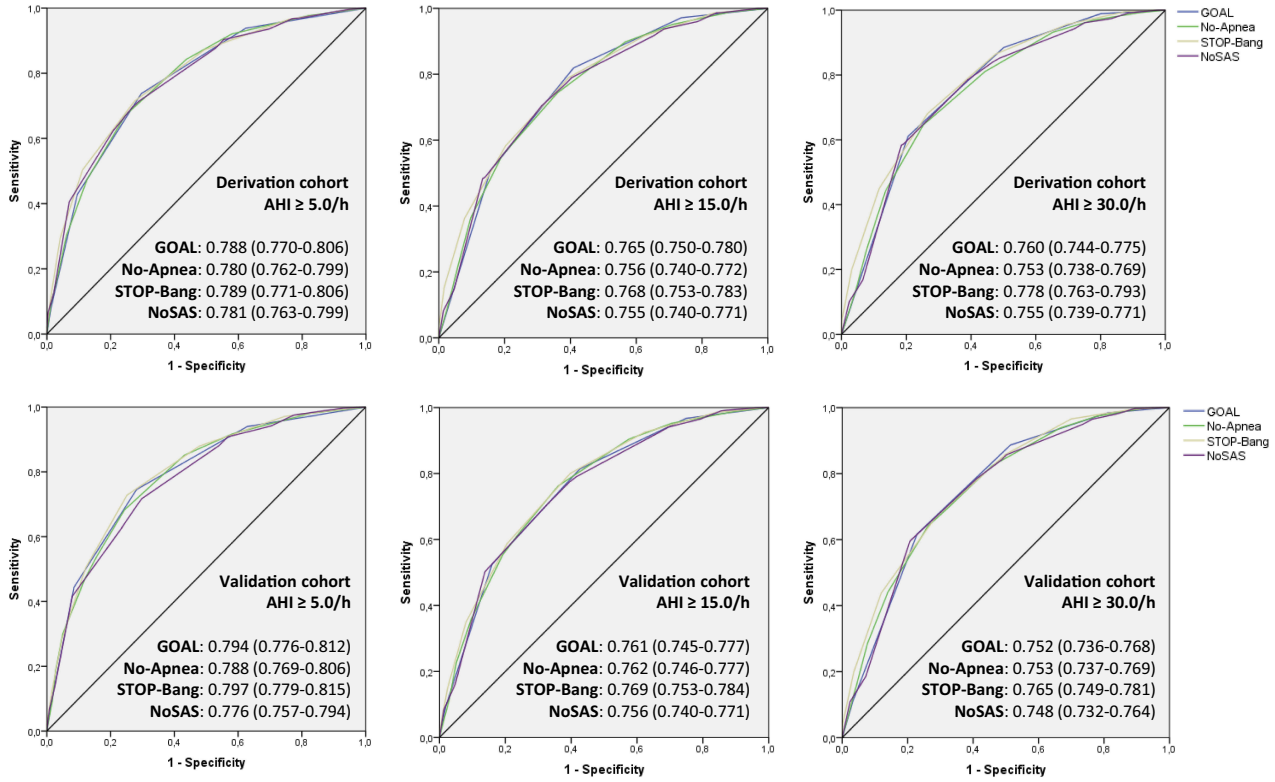

Figure 5 Discriminatory ability, reported as area under the curve (95% confidence interval), of GOAL questionnaire, No-Apnea score, STOP-Bang questionnaire, and NoSAS score for screening of obstructive sleep apnea (OSA) assessed by an apnea/hypopnea index (AHI) ≥5.0/h ($OSA_{\geq 5}$), ≥15.0/h ($OSA_{\geq 15}$), and ≥30.0/h ($OSA_{\geq 30}$). In the derivation cohort, GOAL questionnaire had similar discrimination to No-Apnea, STOP-Bang, and NoSAS for predicting $OSA_{\geq 5}$ (p-values: 0.476, 0.928, and 0.532; respectively), $OSA_{\geq 15}$ (p-values: 0.404, 0.778, and 0.354; respectively), and $OSA_{\geq 30}$ (p-values: 0.557, 0.123, and 0.674; respectively). In the validation cohort, GOAL questionnaire showed equivalent discrimination to No-Apnea, STOP-Bang, and NoSAS for predicting $OSA_{\geq 5}$ (p-values: 0.593, 0.786, and 0.115; respectively); $OSA_{\geq 15}$ (p-values: 0.927, 0.465, and 0.651; respectively), and $OSA_{\geq 30}$ (p-values: 0.936, 0.293, and 0.749; respectively).

improved allocation of patients into corresponding priorities, thus enabling better prioritization of financial resources. In both cohorts, there was no superiority of one given tool over the other, which shows a possible great practical applicability of the GOAL questionnaire, as it contains only four clinical parameters easily obtained during any evaluation of a patient with suspected OSA. As with any population study, future exploration for other world regions and different clinical settings will be critical for widespread implementation of such simple and concise screening tool.

Disclosure

The authors declare no conflicts of interest.

References

1. Kapur VK, Auckley DH, Chowdhuri S, et al. Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American Academy of Sleep Medicine clinical practice guideline. *J Clin Sleep Med*. 2017;13(3):479–504. doi:10.5664/jcsm.6506
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. doi:10.1183/09031936.00022111
3. Pelletier-Fleury N, Meslier N, Gagnadoux F, et al. Economic arguments for the immediate management of moderate-to-severe obstructive sleep apnoea syndrome. *Eur Respir J*. 2004;23(1):53–60. doi:10.1183/09031936.03.00066903
4. Peppard PE, Young T, Barnet JH, et al. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol*. 2013;177(9):1006–1014. doi:10.1093/aje/kws342
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. doi:10.1016/S2213-2600(15)00043-0
6. Tufik S, Santos-Silva R, Taddei JA, et al. Obstructive sleep apnea syndrome in the Sao Paulo epidemiologic sleep study. *Sleep Med*. 2010;11(5):441–446. doi:10.1016/j.sleep.2009.10.005
7. Benjafield AV, Ayas NT, Eastwood PR, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet Respir Med*. 2019;7(8):687–698. doi:10.1016/S2213-2600(19)30198-5
8. Abrishami A, Khajehdehi A, Chung F. A systematic review of screening questionnaires for obstructive sleep apnea. *Can J Anesth*. 2010;57(5):423–438. doi:10.1007/s12630-010-9280-x
9. Rotenberg B, George C, Sullivan K, et al. Wait times for sleep apnea care in Ontario: a multidisciplinary assessment. *Can Resp J*. 2010;17(4):170–174.

10. 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. doi:10.1097/ALN.0b013e31816d83e4
11. Duarte RLM, Fonseca LBM, Magalhães-da-Silveira FJ, et al. Validation of the STOP-Bang questionnaire as a means of screening for obstructive sleep apnea in adults in Brazil. *J Bras Pneumol*. 2017;43(6):456–463. doi:10.1590/S1806-37562017000000139
12. Miller JN, Kupzyk KA, Zimmerman L, et al. Comparisons of measures used to screen for obstructive sleep apnea in patients referred to a sleep clinic. *Sleep Med*. 2018;51:15–21. doi:10.1016/j.sleep.2018.06.007
13. Chia P, Seet E, Macachor JD, et al. The association of preoperative STOP-Bang scores with postoperative critical care admission. *Anaesthesia*. 2013;68(9):950–952. doi:10.1111/anae.12369
14. Nagappa M, Patra J, Wong J, et al. Association of STOP-Bang questionnaire as a screening tool for sleep apnea and postoperative complications: a systematic review and Bayesian meta-analysis of prospective and retrospective cohort studies. *Anesth Analg*. 2017;125(4):1301–1308. doi:10.1213/ANE.0000000000002344
15. Silva GE, Vana KD, Goodwin JL, et al. 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. doi:10.5664/JCSM.1308
16. Tan A, Yin JD, Tan LW, et al. Predicting obstructive sleep apnea using the STOP-Bang questionnaire in the general population. *Sleep Med*. 2016;27–28:66–71. doi:10.1016/j.sleep.2016.06.034
17. 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. doi:10.1016/S2213-2600(16)30075-3
18. Tan A, Hong Y, Tan LWL, et al. Validation of NoSAS score for screening of sleep-disordered breathing in a multiethnic Asian population. *Sleep Breath*. 2017;21(4):1033–1038. doi:10.1007/s11325-016-1455-4
19. Hong C, Chen R, Qing S, et al. Validation of the NoSAS score for the screening of sleep-disordered breathing: a hospital-based retrospective study in China. *J Clin Sleep Med*. 2018;14(2):191–197. doi:10.5664/jcsm.6930
20. Peng M, Chen R, Cheng J, et al. Application value of the NoSAS score for screening sleep-disordered breathing. *J Thorac Dis*. 2018;10(8):4774–4781. doi:10.21037/jtd.2018.07.46
21. Guichard K, Marti-Soler H, Micoulaud-Franchi JA, et al. The NoSAS score: a new and simple screening tool for obstructive sleep apnea syndrome in depressive disorder. *J Affect Disord*. 2018;227:136–140. doi:10.1016/j.jad.2017.10.015
22. Coutinho Costa J, Rebelo-Marques A, Machado JN, et al. Validation of NoSAS (neck, obesity, snoring, age, sex) score as a screening tool for obstructive sleep apnea: analysis in a sleep clinic. *Pulmonology*. 2019;25(5):263–270. doi:10.1016/j.pulmoe.2019.04.004
23. Duarte RL, Rabahi MF, Magalhães-da-Silveira FJ, et al. 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. doi:10.5664/jcsm.7202
24. Duarte RLM, Rabahi MF, Oliveira-e-Sá TS, et al. Fractional exhaled nitric oxide measurements and screening of obstructive sleep apnea in a sleep-laboratory setting: a cross-sectional study. *Lung*. 2019;197(2):131–137. doi:10.1007/s00408-018-0190-y
25. Duarte RLM, Mello FCQ, Magalhães-da-Silveira FJ, et al. Comparative performance of screening instruments for obstructive sleep apnea in morbidly obese patients referred to a sleep laboratory: a prospective cross-sectional study. *Sleep Breath*. 2019;23(4):1123–1132. doi:10.1007/s11325-019-01791-w
26. Duarte RLM, Magalhães-da-Silveira FJ, Oliveira-e-Sá TS, et al. Predicting obstructive sleep apnea in patients with insomnia: a comparative study with four screening instruments. *Lung*. 2019;197(4):451–458. doi:10.1007/s00408-019-00232-5
27. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540–545. doi:10.1093/sleep/14.6.540
28. 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. doi:10.5664/jcsm.2172
29. 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. doi:10.1148/radiology.148.3.6878708
30. Senaratna CV, Perret JL, Matheson MC, et al. Validity of the Berlin questionnaire in detecting obstructive sleep apnea: a systematic review and meta-analysis. *Sleep Med Rev*. 2007;36:116–124. doi:10.1016/j.smrv.2017.04.001
31. Chung F, Subramanyam R, Liao P, et al. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth*. 2012;108(5):768–775. doi:10.1093/bja/aes022
32. 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. doi:10.1007/s11325-017-1482-9
33. Nigro CA, Dibur E, Borsini E, et al. The influence of gender on symptoms associated with obstructive sleep apnea. *Sleep Breath*. 2018;22(3):683–693. doi:10.1007/s11325-017-1612-4
34. Dancey DR, Hanly PJ, Soong C, et al. Gender differences in sleep apnea: the role of neck circumference. *Chest*. 2003;123(5):1544–1550. doi:10.1378/chest.123.5.1544
35. Fietze I, Laharnar N, Obst A, et al. Prevalence and association analysis of obstructive sleep apnea with gender and age differences - Results of SHIP-Trend. *J Sleep Res*. 2019;28(5):e12770. doi:10.1111/jsr.12770
36. Maimon N, Hanly PJ. Does snoring intensity correlate with the severity of obstructive sleep apnea? *J Clin Sleep Med*. 2010;6(5):475–478.
37. Fonseca LB, Silveira EA, Lima NM, et al. STOP-Bang questionnaire: translation to Portuguese and cross-cultural adaptation for use in Brazil. *J Bras Pneumol*. 2016;42(4):266–272. doi:10.1590/S1806-37562015000000243
38. Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. *Anesthesiology*. 2009;110(4):928–939. doi:10.1097/ALN.0b013e31819c47b6