SLEEP BREATHING PHYSIOLOGY AND DISORDERS • ORIGINAL ARTICLE

Comparative performance of screening instruments for obstructive sleep apnea in morbidly obese patients referred to a sleep laboratory: a prospective cross-sectional study

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

Abstract

Purpose Obstructive sleep apnea (OSA) is very common occurrence among morbidly obese patients. Our main objectives were to validate the No-Apnea, a 2-item screening tool, in morbidly obese patients and compare its performance with three other instruments: STOP-Bang questionnaire, NoSAS score, and Epworth Sleepiness Scale (ESS).

Methods A cross-sectional analysis of morbidly obese patients (body mass index [BMI] ≥ 35.0 kg/m^2) grouped into two independent samples: bariatric surgery patients (BS) and non-bariatric surgery patients (NBS). All patients underwent overnight polysomnography. Discriminatory ability was assessed by area under the curve (AUC). OSA severity was defined by apnea/hypopnea index cut-off points: ≥ 5.0/h (OSA$_{\geq 5}$), ≥ 15.0/h (OSA$_{\geq 15}$), and ≥ 30.0/h (OSA$_{\geq 30}$).

Results A total of 1017 subjects (40.4% in BS cohort and 59.6% in NBS cohort) were evaluated. In the BS cohort, No-Apnea had similar discrimination to STOP-Bang and NoSAS for predicting OSA$_{\geq 5}$ ($p = 0.979$ and $p = 0.358$, respectively), OSA$_{\geq 15}$ ($p = 0.158$ and $p = 0.399$, respectively), and OSA$_{\geq 30}$ ($p = 0.388$ and $p = 0.903$, respectively). In the NBS cohort, No-Apnea had similar discrimination to STOP-Bang and NoSAS for predicting OSA$_{\geq 5}$ ($p = 0.528$ and $p = 0.428$, respectively), OSA$_{\geq 15}$ ($p = 0.825$ and $p = 0.108$, respectively), and OSA$_{\geq 30}$ ($p = 0.458$ and $p = 0.186$, respectively). Moreover, No-Apnea performed significantly better than ESS in both BS and NBS cohorts ($p < 0.001$).

Conclusions No-Apnea is a useful and practical tool for screening of OSA in morbidly obese patients, with non-inferior performance to STOP-Bang questionnaire and NoSAS score.

Keywords Obstructive sleep apnea · Polysomnography · Screening · Obesity

✉ Ricardo L. M. Duarte
rlmduarte@gmail.com

1 Sleep - Laboratório de Estudo dos Distúrbios do Sono, Centro Médico BarraShopping, Avenida das Américas 4666, sala 309, Barra da Tijuca, Rio de Janeiro 22649-900, 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 Faculdade de Medicina, Universidade Federal de Goiás, Goiás, Brazil

6 Department of Child Health, University of Missouri School of Medicine, Columbia, MO, USA

Introduction

Obstructive sleep apnea (OSA) is a highly prevalent chronic medical condition [1, 2] and of the several risk factors attributable to its pathophysiology, obesity is by far the more robust predictor [3]. Obesity commonly leads to increases in neck circumference (NC), which can reduce upper airway caliber and facilitate the emergence of increased upper airway collapsibility due to local adipose tissue deposition [4]. Consequently, since obesity is an important risk factor for OSA, a high prevalence of OSA in obese subjects and especially among morbidly obese individuals is to be expected [5, 6]. Similar to what occurs in the general population, OSA is also underdiagnosed in the bariatric population [7, 8], prompting most, but not all, bariatric surgery programs to engage in routine evaluation of all patients for the presence of OSA, regardless of any sleep complaints being reported [9,

10]. Furthermore, bariatric patients with OSA appear to be at high risk for both perioperative and postoperative complications, thereby justifying early recognition of OSA for improved outcomes [11, 12]. Indeed, bariatric patients who did not receive continuous positive airway pressure (CPAP) treatment postoperatively were more likely to develop pulmonary complications [11, 12].

OSA is characterized by recurrent upper airway obstructive episodes, resulting in intermittent hypoxemia, arousals, and has been associated with cardiovascular and metabolic consequences, while also promoting increased overall mortality [13]. Currently, the gold standard for OSA diagnosis consists of overnight polysomnography (PSG) in the laboratory; however, it is not readily available for the large number of patients with suspected OSA, mainly in areas with limited healthcare resources. Using clinical instruments, it is possible to identify patients at high risk for OSA and offer them portable diagnostic methods, and accordingly reduce the long waiting times found in many sleep laboratories [3]. This strategy has already been tested in bariatric patients subjected to unattended diagnostic methods, an approach that apparently demonstrated good diagnostic accuracy [14].

The recently developed and validated instrument "No-Apnea" [15] is a very practical tool with only two objective parameters: NC and age, with a final score ranging from 0 to 9 points. A cut-off point ≥ 3 was previously employed to classify patients at high risk of OSA, with robust areas under the curve (AUCs) for screening of OSA at different levels of severity. Furthermore, despite its simplicity, when compared to two previously validated and widely used instruments (i.e., STOP-Bang questionnaire [16] and NoSAS score [17]), there were no statistically significant differences in performance between No-Apnea and the other tools [15].

Notwithstanding, there is only a scarce number of studies using questionnaires in the morbidly obese population [5, 18, 19], such that we hypothesized that screening of OSA risk with No-Apnea would provide a reliable approach in morbidly obese patients. For the present study, our main objectives were to evaluate the predictive performance of No-Apnea in a large sample of morbidly obese Brazilian patients referred for PSG evaluation in the context of bariatric surgery (BS cohort) or referred for clinical suspicion of OSA (NBS cohort), and also to compare No-Apnea discriminatory performance with three other instruments, namely STOP-Bang [16], NoSAS [17], and Epworth Sleepiness Scale (ESS) [20].

Methods

Study design

This cross-sectional study, from January 2017 to August 2018, prospectively enrolled morbidly obese subjects who were clinically referred for overnight in-laboratory PSG evaluation by their respective treating physicians. Inclusion criteria consisted of age ≥ 18 years and body mass index (BMI) ≥ 35.0 kg/m^2. Exclusion criteria were as follows: previously diagnosed OSA, use of portable sleep studies, incomplete clinical data, and technically inadequate PSG. All morbidly obese subjects enrolled were grouped into two independent groups: BS (patients undergoing assessment prior to bariatric surgery) and NBS (morbidly obese patients referred for suspected OSA). The study protocol was approved by the Ethics Committee of the Federal University of Rio de Janeiro (no. 1.764.165) and was carried out in accordance with the Declaration of Helsinki. Written informed consent was obtained from each subject and anonymity of each participant was preserved.

On the evening of the PSG, clinical data were collected in all patients: gender, age, BMI, NC, self-reported comorbidities (smoking, diabetes mellitus, and hypertension), and data from each of the screening instruments (No-Apnea, STOP-Bang, NoSAS, and ESS). The BMI was calculated by dividing the weight in kilograms by the square of the height in meters (kg/m^2), while NC (in cm) was systematically measured using a flexible tape with all subjects in the upright sitting position, with the upper edge of the tape measure placed immediately below the laryngeal prominence and applied perpendicularly to the long axis of the neck.

Screening tools

The No-Apnea scoring system [15] contains two objective parameters: NC and age. 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), while age is scored as follows: 35–44 years (1 point), 45–54 years (2 points), ≥ 55 years (3 points). The points for each variable are added reaching a final score of 0–9 points, being a cut-off point ≥ 3 used to classify patients at high risk of OSA [15].

The STOP-Bang questionnaire [16] was originally developed for the screening of OSA in preoperative patients and consists of eight yes-or-no questions: loud snoring, tiredness, observed apnea, hypertension, BMI > 35 kg/m^2, age > 50 years, NC > 40 cm, and male gender. A STOP-Bang score ≥ 3 points is considered as positive; however, a cut-off point ≥ 4 is more suitable to screen obese and morbidly obese subjects [21], being that cut-off point ≥ 4 was employed in our study.

The NoSAS score [17] is a recently reported screening tool containing five parameters: 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 gender (2 points), totaling a score of 0–17 points. A score ≥ 8 points was considered as high risk for OSA [17].

The ESS [20] is an 8-item questionnaire that assesses the subjective likelihood of falling asleep in various contexts: each item is scored from zero (never sleeps) to three (high

chance of falling asleep), with a final score from 0 to 24 points. A score ≥ 11 points was considered as excessive daytime somnolence [20].

Sleep studies

All tests were conducted in *Sleep Laboratory at Centro Medico BarraShopping*, Rio de Janeiro, Brazil. All patients underwent an attended, in-lab full PSG (EMBLA® S7000, Embla Systems, Inc., Broomfield, CO, USA) with video monitoring during the whole night, consisting of recording of electroencephalography, electrooculography, electromyography (chin and legs), electrocardiography, airflow, thoracic and abdominal impedance belts, oxygen saturation (SpO_2), microphone for snoring, and sensors for body position. Polysomnographic records were scored manually in accordance with the latest 2012 American Academy of Sleep Medicine (AASM) guidelines [22] by two board-certified sleep physicians, both of whom were blinded for No-Apnea, STOP-Bang, NoSAS, and ESS results. These data included total sleep time (TST), sleep efficiency, sleep stages, rapid-eye movement (REM) and sleep latencies, arousal index, apnea/hypopnea index (AHI), and SpO_2 values. Apneas were defined as a decrease of at least 90% of airflow from baseline, lasting at least 10 s, while hypopneas were classified with a decrease $\geq 30\%$ of pre-event during ≥ 10 s associated with $\geq 3\%$ oxygen desaturation or an arousal [22]. The AHI was calculated as the sum of the number of apneas and hypopneas per hour of sleep, with OSA severity assessed by three different AHI cut-off values: ≥ 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; Chicago, IL, USA). Results are shown as mean $\pm$ SD for quantitative variables and as number (n) and percentage (%) for qualitative variables. Comparisons between groups were performed using the chi-square test for dichotomous variables, Student's t test, and univariate analysis of variance (ANOVA) for continuous variables. Correlation was evaluated by Spearman correlation coefficient (r). Discrimination was estimated from the AUC obtained by receiver operator characteristic (ROC) curves. The AUCs obtained were compared using a validated algorithm [23]. Calibration was evaluated by Hosmer-Lemeshow chi-square test ($p < 0.05$ considered as poor calibration). Overall performance was evaluated using the Nagelkerke R^2. Using the 2×2 contingency tables, the following parameters were calculated: sensitivity, specificity, positive and negative predictive values (PPV and NPV, respectively), and accuracy (rate of correctly screened patients). All estimates were reported with their respective 95% confidence intervals (CIs). A two-tailed p value < 0.05 was considered statistically significant.

Results

The flowchart of the study is reported in Fig. 1. Overall, 1017 morbidly obese subjects were allocated into two independent cohorts: BS ($n = 411$; 40.4%) and NBS ($n = 606$; 59.6%). As shown in Table 1, BS patients (vs. NBS patients) were younger, had a predominance of women, and had a higher BMI (all with $p < 0.001$). All polysomnographic measures (Table 1) were statistically different between BS and NBS groups, except for TST ($p = 0.139$), sleep efficiency ($p = 0.333$), stage N1 sleep ($p = 0.148$), stage REM sleep ($p = 0.072$), and hypopnea index ($p = 0.419$). Prevalence of OSA was higher in the NBS than in the BS cohort: (i) $OSA_{\geq 5}$: 92.4 vs. 84.7% ($p < 0.001$), (ii) $OSA_{\geq 15}$: 77.2 vs. 62.5% ($p < 0.001$), and (iii) $OSA_{\geq 30}$: 59.1 vs. 41.4% ($p < 0.001$). In the BS cohort, the prevalence of OSA was higher in men than in women: (i) $OSA_{\geq 5}$: 97.1 vs. 78.3% ($p < 0.001$), $OSA_{\geq 15}$: 87.8 vs. 49.6% ($p < 0.001$), and $OSA_{\geq 30}$: 73.4 vs. 25.0% ($p < 0.001$). Similarly, in the NBS cohort, the prevalence of OSA was higher in men than in women: (i) $OSA_{\geq 5}$: 96.3 vs. 87.0% ($p < 0.001$), (ii) $OSA_{\geq 15}$: 86.1 vs. 64.8% ($p < 0.001$), and (iii) $OSA_{\geq 30}$: 75.6 vs. 36.0% ($p < 0.001$). Figure 2 illustrates that the frequency of patients assessed as high risk was always higher in the NBS cohort than in the BS cohort, for all screening instruments evaluated.

Characteristics of the No-Apnea

The mean No-Apnea score was significantly lower in BS patients compared to NBS patients (4.7 ± 2.5 points vs. 5.7 ± 2.3 points; $p < 0.001$). Similarly, mean AHI values increased linearly according to increasing No-Apnea scores (from 0 to 9 points): in the BS cohort, the mean AHI ranged from 9.7 ± 11.1 to 69.8 ± 22.9/h (p value for trend < 0.001), while in the NBS cohort, it ranged from 14.1 ± 21.8 to 49.1 ± 28.7/h (p value for trend < 0.001). As can be seen in Table 2, No-Apnea showed statistically significant correlation with all respiratory measures during sleep obtained in the PSG (all with $p < 0.001$). According to Table 3, the No-Apnea instrument performed similarly among genders for screening of $OSA_{\geq 5}$ ($p = 0.973$) and $OSA_{\geq 15}$ ($p = 0.817$), while it performed significantly better in women than in men for screening of $OSA_{\geq 30}$ ($p = 0.033$).

In the BS cohort, the No-Apnea instrument showed adequate calibration and overall performance for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$: Hosmer-Lemeshow test 4.446 ($p = 0.727$) and Nagelkerke R^2 0.272; Hosmer-Lemeshow test

```
┌─────────────────────────────────┐
│   All morbidly obese patients   │
│   (age ≥ 18 years) referred     │
│   for sleep-lab                 │
│        (n = 1,170)              │
└─────────────────────────────────┘
              │
              │         ┌──────────────────────────────────────────┐
              │────────▶│  153 patients were excluded:             │
              │         │                                          │
              │         │  Incomplete data and/or inadequate PSG   │
              │         │  (n = 74)                                │
              │         │  Diagnosis with portable monitors (n=60) │
              │         │  OSA previously diagnosed (n = 10)       │
              │         │  Refuse to sign the consent form (n = 9) │
              │         └──────────────────────────────────────────┘
              ▼
┌─────────────────────────────────┐
│     All eligible patients       │
│         (n = 1,017)             │
└─────────────────────────────────┘
         │              │
         ▼              ▼
┌──────────────┐  ┌──────────────┐
│  Patients    │  │  Patients    │
│ allocated    │  │ allocated    │
│ into BS      │  │ into NBS     │
│ cohort       │  │ cohort       │
│ (n = 411)    │  │ (n = 606)    │
└──────────────┘  └──────────────┘
```
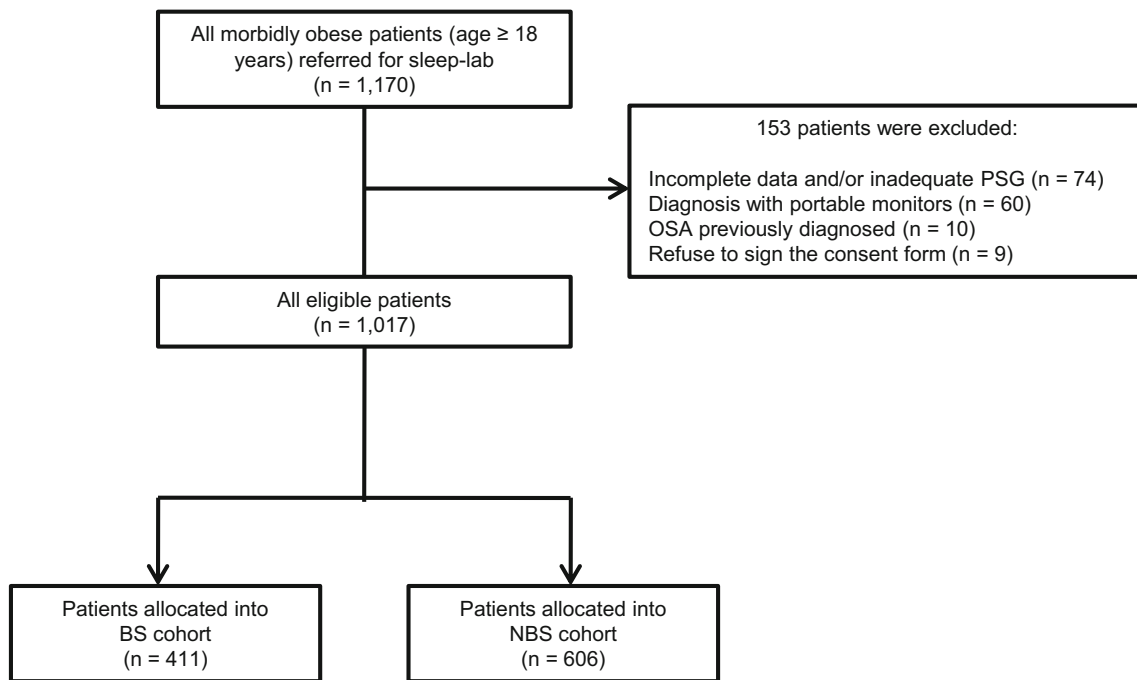
Fig. 1 The flowchart of the patients. OSA obstructive sleep apnea, PSG polysomnography, BS bariatric surgery, and NBS non-bariatric surgery

5.337 ($p = 0.619$) and Nagelkerke R^2 0.351; and Hosmer-Lemeshow test 1.321 ($p = 0.995$) and Nagelkerke R^2 0.369; respectively. Similarly, in the NBS cohort, our model displayed adequate calibration and overall performance for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$: Hosmer-Lemeshow test 5.266 ($p = 0.729$) and Nagelkerke R^2 0.236; Hosmer-Lemeshow test 9.583 ($p = 0.143$) and Nagelkerke R^2 0.197; and Hosmer-Lemeshow test 13.237 ($p = 0.067$) and Nagelkerke R^2 0.223; respectively.

Performance of all instruments

Table 4 summarizes the predictive performance of all instruments evaluated. In the BS cohort, for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$, No-Apnea showed sensitivity ranging from 80.7 to 93.5% and specificity ranging from 42.9 to 34.4%, respectively. In addition, No-Apnea showed the highest sensitivity among all instruments: 80.7% ($OSA_{\geq 5}$), 87.9% ($OSA_{\geq 15}$), and 93.5% ($OSA_{\geq 30}$). In the NBS cohort, No-Apnea reported the following characteristics for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$: sensitivity ranging from 91.6 to 96.9%, specificity ranging from 47.8 to 23.4%, and accuracy ranging from 88.3 to 66.8%; respectively. As shown in Fig. 3, No-Apnea had non-inferior discrimination to STOP-Bang and NoSAS for screening of $OSA_{\geq 5}$, $OSA_{\geq 15}$, and $OSA_{\geq 30}$. The ESS was not adequate in the prediction of OSA, and No-Apnea performed significantly better than ESS in both of the cohorts evaluated ($p < 0.001$).

Discussion

The present study showed that No-Apnea—an extremely simple, practical, and objective instrument containing only two parameters—can be used satisfactorily as a screening tool for OSA among morbidly obese patients, regardless of whether the indication for PSG was clinical suspicion of OSA or preoperative evaluation before bariatric surgery. No-Apnea showed adequate discrimination, calibration, and overall performance in both morbidly obese groups being evaluated. Moreover, despite its apparent simplicity, its discriminatory ability was not inferior to that achieved by the STOP-Bang or NoSAS, at all levels of OSA severity. The simplicity and ease aspects of No-Apnea can obviously confer a significant advantage, since STOP-Bang and NoSAS contain more parameters (eight and five parameters, respectively) without any incremental value being added. Moreover, No-Apnea is exclusively composed of objective parameters, thereby reducing the subjective information bias, and can also be used in individuals who sleep alone, where subjective sleep information from a bed partner is not available. The simplicity of No-Apnea can facilitate its widespread implementation as a screening model for morbidly obese individuals, allowing the referral of these patients towards diagnosis using home-based portable diagnostic systems, thus avoiding the long waiting times for full PSG. Of note, similar findings and conclusions emerged during study of development and validation of the No-Apnea [15].

As anticipated, we found a high prevalence of OSA in this large sample of consecutively enrolled morbidly

Table 1 General and sleep characteristics (n = 1017)

Parameter	BS cohort (n = 411)	NBS cohort (n = 606)	p value
Clinical data			
Female gender	272 (66.2)	253 (41.7)	< 0.001
Age, years	39.0 ± 10.9	45.3 ± 13.1	< 0.001
BMI, kg/m^2	42.9 ± 5.3	39.0 ± 3.9	< 0.001
NC, cm	42.7 ± 4.9	43.3 ± 4.7	0.033
Current smokers	38 (9.2)	51 (8.4)	0.653
Hypertension	187 (45.5)	302 (49.8)	0.180
Diabetes mellitus	67 (16.3)	114 (18.8)	0.317
Screening tools			
No-Apnea, points	4.7 ± 2.5	5.7 ± 2.3	< 0.001
STOP-Bang, points	4.1 ± 1.5	5.3 ± 1.4	< 0.001
NoSAS, points	10.7 ± 2.7	12.1 ± 2.7	< 0.001
ESS, points	9.5 ± 4.5	10.7 ± 5.1	< 0.001
Polysomnographic data			
Total sleep time, min	331.7 ± 71.3	338.3 ± 69.4	0.139
Sleep efficiency, %	77.2 ± 15.2	78.1 ± 15.0	0.333
Sleep latency, min	43.9 ± 42.0	37.2 ± 40.0	0.011
REM sleep latency, min	146.9 ± 78.2	157.8 ± 83.8	0.045
Stage N1 sleep, %	5.0 ± 5.4	4.5 ± 5.2	0.148
Stage N2 sleep, %	67.7 ± 12.0	71.5 ± 12.5	< 0.001
Stage N3 sleep, %	11.2 ± 8.3	8.9 ± 8.3	< 0.001
Stage REM sleep, %	16.0 ± 8.2	15.0 ± 8.2	0.072
Arousal index, n/h	35.7 ± 29.7	43.5 ± 29.5	< 0.001
AHI, n/h	33.9 ± 31.9	43.3 ± 31.3	< 0.001
AI, n/h	14.5 ± 26.3	23.1 ± 29.8	< 0.001
HI, n/h	19.3 ± 17.0	20.1 ± 15.2	0.419
Mean SpO$_2$, %	93.5 ± 2.7	92.3 ± 3.5	< 0.001
Lowest SpO$_2$, %	81.2 ± 8.6	77.7 ± 9.9	< 0.001
ODI at 3%, n/h	29.0 ± 30.2	38.0 ± 30.3	< 0.001

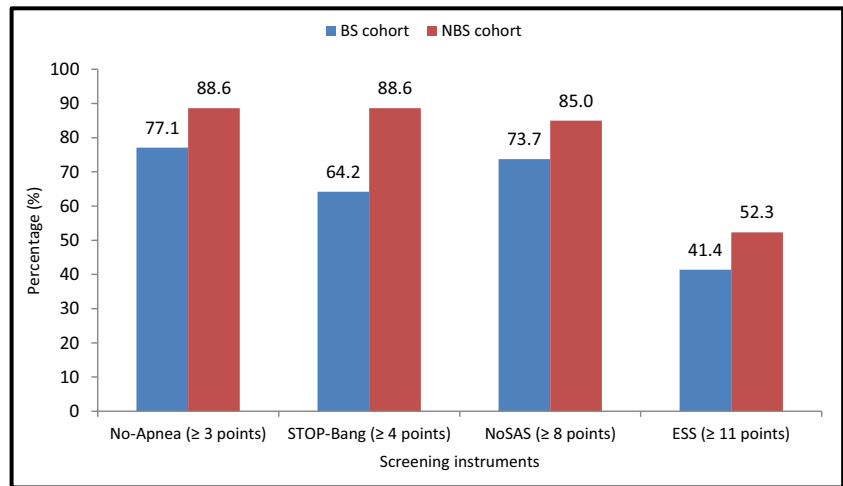Data are presented as means ± SD or n (%)

BS bariatric surgery, *NBS* non-bariatric surgery, *BMI* body-mass index, *NC* neck circumference, *ESS* Epworth Sleepiness Scale, *REM* rapid-eye movement, *AHI* apnea/hypopnea index, *AI* apnea index, *HI* hypopnea index, *SpO$_2$* oxygen saturation, *ODI* oxygen desaturation index

obese patients. Previous studies have also reported a high prevalence of OSA in BS patients [5, 7–9]. Additionally, BS patients were younger, more obese, and had a predominance of women when compared to NBS patients, a finding that is compatible with other studies in bariatric cohorts [5, 9, 10], but not in studies of the general population [1, 2] or in sleep-lab setting [24], in which there is a well-established male predominance. In addition, NBS patients had increased severity of sleep-disordered breathing than BS patients reflecting the compulsory referral of BS subjects independently from the presence OSA-evoking symptoms. We are unaware of any studies to date comparing clinical and polysomnographic data in two distinct groups of morbidly obese patients according to their indication for PSG.

As reported above, our findings showed that the two cohorts that were evaluated (BS and NBS cohorts) presented clear differences in their demographic and anthropometric characteristics, as illustrated by a higher prevalence of OSA$_{\geq 5}$, OSA$_{\geq 15}$, and OSA$_{\geq 30}$ in the NBS cohort compared to the BS cohort. The differences in the prevalence of OSA may account for the performance differences of No-Apnea in the BS and NBS groups, whereby it performed better in the bariatric group, particularly in the more severe categories of OSA (OSA$_{\geq 15}$ and OSA$_{\geq 30}$).

The STOP-Bang questionnaire [16], a widely used screening tool, was initially developed in surgical patients to screen for the presence of OSA and showed a sensitivity 83.6%, specificity 56.4%, PPV 81.0%, and NPV 60.8%. Although the original STOP-Bang [16] uses a cut-off point ≥ 3 to

Fig. 2 Percentage (%) of patients at high risk for diagnosis of obstructive sleep apnea according to four screening instruments evaluated: No-Apnea, STOP-Bang, NoSAS, and Epworth Sleepiness Scale (ESS). BS bariatric surgery and NBS non-bariatric surgery

indicate patients with high risk of OSA, when applied to obese subjects, a cut-off point ≥ 4 is apparently more suitable [21], as illustrated by a high sensitivity (87.5%) and high NPV (90.5%), while in morbidly obese patients, the same score showed a sensitivity of 89.5% and a specificity of 25.5% [21]. The favorable yield of the STOP-Bang in screening patients for OSA among different populations was corroborated in a recent meta-analysis [25].

However, in a single-center retrospective study with 266 bariatric patients evaluated with standard overnight PSG, it was observed that neither STOP-Bang nor the Berlin questionnaire were useful screening models in this surgical population [26]. Interestingly, the results of this study were substantially different from others, being the performance obtained by the screening models substantially lower than previously reported in the literature [26]. In addition, ESS was also not predictive of $OSA_{\geq 30}$

(AUC 0.557; 95% CI 0.476–0.639) or $OSA_{\geq 15}$ (AUC 0.512; 95% CI 0.440–0.584) [26].

The recently developed NoSAS score [17] performed well in a general population sample from Switzerland included in the HypnoLaus cohort with an AUC of 0.74. In another general population cohort from Brazil, the EPISONO cohort, the NoSAS showed an AUC of 0.81. The NoSAS score performed significantly better than STOP-Bang questionnaire ($p < 0.0001$) and Berlin questionnaire ($p < 0.0001$) [17]. However, in a study [27] of a multiethnic Asian cohort, STOP-Bang, Berlin questionnaire, and NoSAS score performed similarly, with AUCs being clustered around 0.682–0.748. In another study [28] aiming to validate the NoSAS as a screening tool for OSA in clinical populations, this tool showed an AUC of 0.707, performing significantly better than the STOP (AUC: 0.655), STOP-Bang (AUC: 0.704), and the ESS (AUC: 0.642).

A previous study applied four models (ESS, Fatigue Severity Scale, STOP-Bang, and NoSAS) to 251 bariatric patients (76% females): STOP-Bang and NoSAS performed better than the ESS and Fatigue Severity Scale. Except for the ESS, all sleep questionnaires allowed better OSA prediction in women than in men [19]. However,

Table 2 Spearman correlation between No-Apnea scores and respiratory parameters obtained by polysomnography ($n = 1017$)

Parameter	No-Apnea scores			
	BS cohort ($n = 411$)		NBS cohort ($n = 606$)	
	Correlation (r)	p value	Correlation (r)	p value
Polysomnographic data				
AHI, n/h	0.566	< 0.001	0.406	< 0.001
AI, n/h	0.507	< 0.001	0.395	< 0.001
HI, n/h	0.402	< 0.001	0.165	< 0.001
Mean SpO$_2$, %	− 0.438	< 0.001	− 0.370	< 0.001
Lowest SpO$_2$, %	− 0.473	< 0.001	− 0.273	< 0.001
ODI at 3%, n/h	0.559	< 0.001	0.372	< 0.001

BS bariatric surgery, *NBS* non-bariatric surgery, *AHI* apnea/hypopnea index, *AI* apnea index, *HI* hypopnea index, *SpO₂* oxygen saturation, *ODI* oxygen desaturation index

Table 3 No-Apnea discriminatory ability according to gender ($n = 1017$)

	Females ($n = 525$)	Males ($n = 492$)	p value
AHI ≥ 5.0/h	0.719 (0.662–0.776)	0.717 (0.589–0.846)	0.973
AHI ≥ 15.0/h	0.689 (0.643–0.734)	0.680 (0.606–0.754)	0.817
AHI ≥ 30.0/h	0.702 (0.655–0.749)	0.621 (0.561–0.681)	0.033

Data are presented as estimates (95% confidence intervals)

AHI apnea/hypopnea index

Table 4 Predicting obstructive sleep apnea with No-Apnea, STOP-Bang, NoSAS, and Epworth Sleepiness Scale (ESS)

	No-Apnea	STOP-Bang	NoSAS	ESS
BS cohort (n = 411)				
AHI ≥ 5.0/h				
Sensitivity	80.7 (78.7–82.8)	71.3 (69.1–73.0)	77.3 (75.2–79.4)	41.4 (39.2–43.4)
Specificity	42.9 (31.7–54.4)	74.6 (62.9–83.9)	46.0 (34.6–57.7)	58.7 (46.6–70.0)
PPV	88.6 (86.4–90.9)	93.9 (91.1–96.2)	88.8 (86.4–91.2)	84.7 (80.2–88.9)
NPV	28.7 (21.3–36.4)	32.0 (27.0–36.0)	26.9 (20.2–33.6)	15.4 (12.2–18.3)
Accuracy	74.9 (71.5–78.5)	71.8 (68.2–74.6)	72.5 (69.0–76.1)	44.0 (40.3–47.5)
AHI ≥ 15.0/h				
Sensitivity	87.9 (84.6–90.9)	77.8 (74.1–81.3)	84.8 (81.3–88.1)	42.0 (38.1–45.8)
Specificity	40.9 (35.3–45.9)	58.4 (52.2–64.2)	44.8 (38.9–50.2)	59.7 (53.2–66.1)
PPV	71.3 (68.6–73.7)	75.8 (72.1–79.1)	71.9 (69.0–74.7)	63.5 (57.6–69.3)
NPV	67.0 (57.9–75.2)	61.2 (54.7–67.3)	63.9 (55.5–71.6)	38.2 (34.0–42.2)
Accuracy	70.3 (66.1–74.1)	70.6 (65.9–74.9)	69.8 (65.4–73.9)	48.7 (43.8–53.4)
AHI ≥ 30.0/h				
Sensitivity	93.5 (89.2–96.5)	84.1 (78.8–88.6)	91.2 (86.5–94.6)	41.2 (35.3–47.1)
Specificity	34.4 (31.4–36.5)	49.8 (46.0–53.0)	38.6 (35.3–41.0)	58.5 (54.4–62.7)
PPV	50.2 (47.8–51.7)	54.2 (50.7–57.1)	51.2 (48.5–53.1)	41.2 (35.3–47.1)
NPV	88.3 (80.4–93.6)	81.6 (75.4–86.8)	86.1 (78.7–91.6)	58.5 (54.4–62.7)
Accuracy	58.9 (55.3–61.3)	64.0 (59.6–67.7)	60.3 (56.5–63.2)	51.3 (46.5–56.3)
NBS cohort (n = 606)				
AHI ≥ 5.0/h				
Sensitivity	91.6 (90.5–92.7)	91.1 (90.0–92.2)	88.0 (86.9–89.2)	53.4 (52.2–54.5)
Specificity	47.8 (34.3–61.3)	41.3 (28.3–55.1)	52.2 (38.1–65.8)	60.9 (45.9–74.2)
PPV	95.5 (94.4–96.7)	95.0 (93.9–96.2)	95.7 (94.5–96.9)	94.3 (92.2–96.3)
NPV	31.9 (22.9–40.9)	27.5 (18.8–36.7)	26.4 (19.3–33.3)	9.7 (7.3–11.8)
Accuracy	88.3 (86.2–90.3)	87.3 (85.3–89.4)	85.3 (83.2–87.4)	54.0 (51.7–56.0)
AHI ≥ 15.0/h				
Sensitivity	93.8 (92.1–95.4)	92.9 (91.3–94.6)	90.4 (88.5–92.2)	54.1 (51.8–56.3)
Specificity	29.0 (23.3–34.3)	26.1 (20.4–31.6)	33.3 (27.1–39.5)	53.6 (45.9–61.2)
PPV	81.8 (80.3–83.1)	81.0 (79.5–82.4)	82.1 (80.5–83.8)	79.8 (76.5–83.1)
NPV	58.0 (46.5–68.7)	52.2 (40.8–63.3)	50.5 (41.0–59.9)	25.6 (21.9–29.2)
Accuracy	79.0 (76.4–81.5)	77.7 (75.1–80.2)	77.4 (74.5–80.2)	54.0 (50.5–57.4)
AHI ≥ 30.0/h				
Sensitivity	96.9 (94.9–98.3)	96.6 (94.6–98.1)	95.3 (92.9–97.0)	55.9 (52.4–59.3)
Specificity	23.4 (20.4–25.4)	23.0 (20.0–25.1)	29.8 (26.5–32.4)	52.8 (47.8–57.7)
PPV	64.6 (63.3–65.5)	64.4 (63.0–65.4)	66.2 (64.6–67.4)	63.1 (59.2–66.9)
NPV	84.1 (73.4–91.3)	82.6 (71.8–90.2)	81.3 (72.2–88.2)	45.3 (41.1–49.5)
Accuracy	66.8 (64.4–68.5)	66.5 (64.0–68.2)	68.5 (65.7–70.6)	54.6 (50.5–58.6)

Data are presented as estimates (95% confidence intervals). No-Apnea is a two-item model: neck circumference (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), while 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. STOP-Bang is an eight-item model (1 point for each positive answer): loud snoring, tiredness, observed apnea, hypertension, body mass index (BMI) > 35 kg/m^2, age > 50 years, NC > 40 cm, and male gender; totaling a score of 0–8 points. NoSAS is a five-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 gender (2 points); totaling a score of 0–17 points. ESS is an eight-item questionnaire that assesses the subjective likelihood of falling asleep in various contexts: each item is scored from zero (never sleeps) to three (high chance of falling asleep), totaling a score of 0–24 points. The cut-off points employed to identify patients at high-risk were ≥ 3 points (No-Apnea), ≥ 4 points (STOP-Bang), ≥ 8 points (NoSAS), and ≥ 11 points (ESS).

AHI apnea/hypopnea index, *PPV* positive predictive value, *NPV* negative predictive value, *BS* bariatric surgery, *NBS* non-bariatric surgery
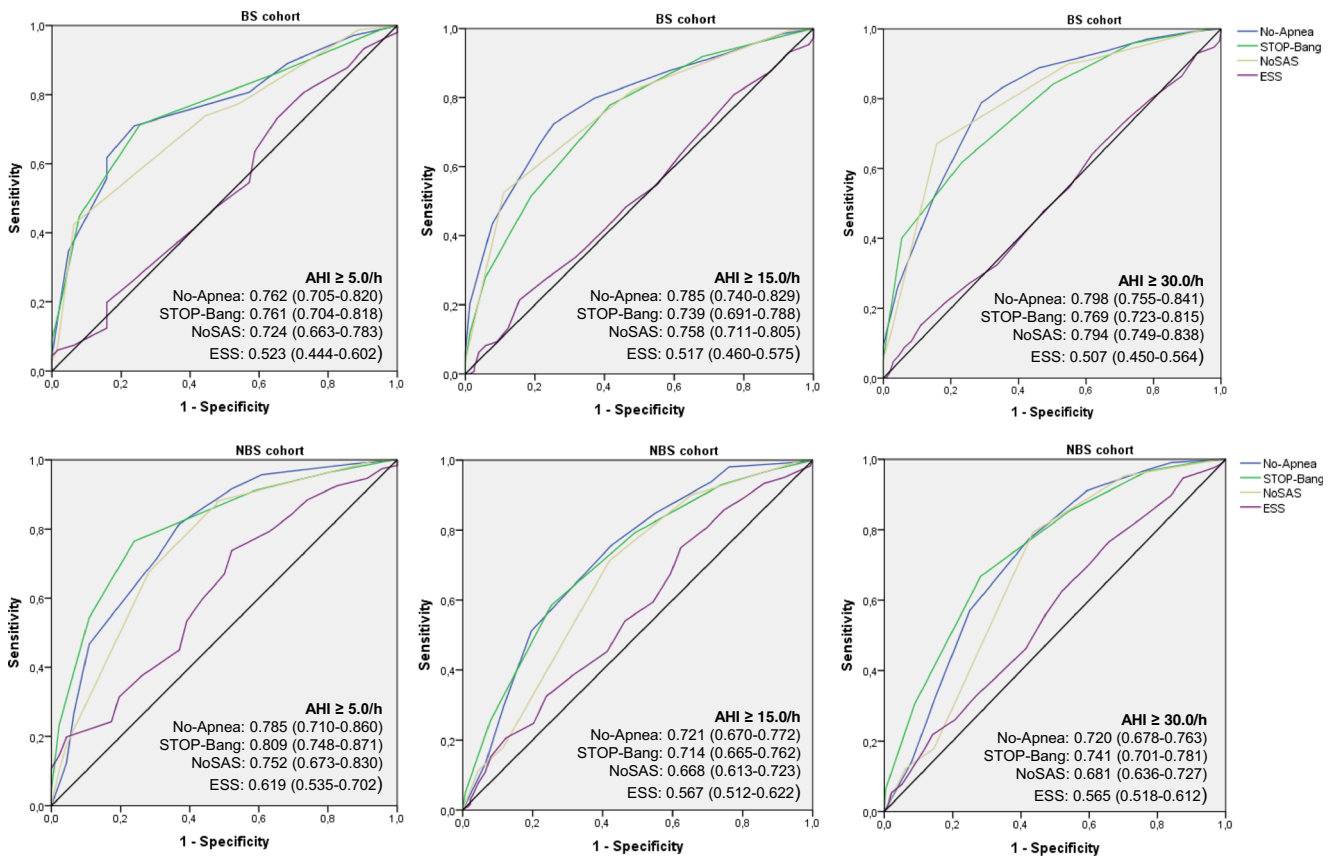
Fig. 3 Discriminatory performance, reported as area under the curve (95% confidence interval), of No-Apnea, STOP-Bang, NoSAS, and Epworth Sleepiness Scale (ESS) for screening of obstructive sleep apnea (OSA) assessed by an apnea/hypopnea index (AHI) $\geq$ 5.0/h (OSA$_{\geq 5}$), $\geq$ 15.0/h (OSA$_{\geq 15}$), and $\geq$ 30.0/h (OSA$_{\geq 30}$). In the bariatric surgery (BS) cohort, No-Apnea had similar discrimination to STOP-Bang and NoSAS for predicting OSA$_{\geq 5}$ ($p = 0.979$ and $p = 0.358$, respectively), OSA$_{\geq 15}$ ($p = 0.158$ and $p = 0.399$, respectively), and OSA$_{\geq 30}$ ($p = 0.388$ and $p = 0.903$, respectively). In the non-bariatric surgery (NBS) cohort, No-Apnea had similar discrimination to STOP-Bang and NoSAS for predicting OSA$_{\geq 5}$ ($p = 0.528$ and $p = 0.428$, respectively), OSA$_{\geq 15}$ ($p = 0.825$ and $p = 0.108$, respectively), and OSA$_{\geq 30}$ ($p = 0.458$ and $p = 0.186$, respectively). The ESS was not a useful screening model for OSA in both subsamples. No-Apnea performed significantly better than ESS in both BS and NBS cohorts ($p < 0.001$)

this study presented some limitations that deserve mention: (i) retrospective study design, (ii) patients were evaluated with polygraphy rather than PSG, and (iii) few men were enrolled in the final analysis ($n = 60$). Although our study was not designed to evaluate gender differences in the performance of No-Apnea among morbidly obese patients, some of the findings deserve mention: there were no gender-related differences in the discriminatory ability for screening of OSA$_{\geq 5}$ and OSA$_{\geq 15}$; however, for screening of OSA$_{\geq 30}$, No-Apnea performed significantly better in women than in men.

This present study also reported that the ESS was not shown to be useful as a screening tool for OSA in morbidly obese patients. Although ESS has been widely used in clinical practice, some studies show a poor utility of ESS as a screening model for OSA [29, 30], possibly because excessive daytime sleepiness is not always present in individuals suffering from OSA.

Strengths and limitations

Our study had some limitations that deserve comment. First, selection of patients occurred in a sleep laboratory, where a high prevalence of OSA is anticipated, and therefore the possibility of selection bias is plausible. In addition, the study was conducted at a single institution, which may limit the reproducibility of our findings in other populations. Despite these limitations, our study enrolled a large sample of consecutive morbidly obese patients. All patients underwent in-lab PSG, which was scored manually by two experienced physicians according to 2012 AASM guidelines [22]. To our knowledge, this is the first study that was effectively designed to assess differences in No-Apnea, STOP-Bang, NoSAS, and ESS predictive performance among morbidly obese subjects grouped into two independent and different groups of obese patients, namely BS and NBS.

Conclusions

OSA is a very common disorder among morbidly obese patients. A very pragmatic and objective screening tool, No-Apnea, showed adequate predictive performance for diagnosis of OSA, and no statistically significant differences emerged when compared to STOP-Bang or NoSAS, at all levels of OSA severity. On the other hand, ESS did not present satisfactory discrimination as OSA screening model in morbidly obese patients.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The authors declare that the protocol of the research work was approved by the Ethics Commission of the Universidade Federal do Rio de Janeiro (UFRJ) according to the Helsinki Declaration, and all the patients provided written informed consent.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM (2013) Increased prevalence of sleep-disordered breathing in adults. Am J Epidemiol 177:1006–1014

2. Heinzer R, Vat S, Marques-Vidal P, Martin-Soler H, Andries D, Tobback N, Mooser V, Preisig M, Malhotra A, Waeber G, Vollenweider P, Tafti M, Haba-Rubio J (2015) Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. Lancet Respir Med 3:310–318

3. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, Harrod CG (2017) Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American Academy of Sleep Medicine clinical practice guideline. J Clin Sleep Med 13:479–504

4. Mortimore I, Marshall I, Wraith PK, Sellar RJ, Douglas NJ (1998) Neck and total body fat deposition in nonobese and obese patients with sleep apnea compared with that in control subjects. Am J Respir Crit Care Med 157:280–283

5. Kositanurit W, Muntham D, Udomsawaengsup S, Chirakalwasan N (2018) Prevalence and associated factors of obstructive sleep apnea in morbidly obese patients undergoing bariatric surgery. Sleep Breath 22:251–256

6. Duarte RL, Magalhães-da-Silveira FJ (2015) 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 41:440–448

7. Rasmussen JJ, Fuller WD, Ali MR (2012) Sleep apnea syndrome is significantly underdiagnosed in bariatric surgical patients. Surg Obes Relat Dis 8:569–573

8. Ravesloot MJ, van Maanen JP, Hilgevoord AA, van Wagensveld BA, de Vries N (2012) Obstructive sleep apnea is underrecognized and underdiagnosed in patients undergoing bariatric surgery. Eur Arch Otorhinolaryngol 269:1865–1871

9. O'Keeffe T, Patterson EJ (2004) Evidence supporting routine polysonography before bariatric surgery. Obes Surg 14:23–26

10. Nepomnayshy D, Hesham W, Erikson B, MacDonald J, Iorio R, Brams D (2013) Sleep apnea: is routine preoperative screening necessary? Obes Surg 23:287–291

11. Kong WT, Chopra S, Kopf M, Morales C, Khan S, Zuccala K, Choi L, Chronakos J (2016) Perioperative risks of untreated obstructive sleep apnea in the bariatric surgery patient: a retrospective study. Obes Surg 26:2886–2890

12. Carron M, Zarantonello F, Tellaroli P, Ori C (2016) Perioperative noninvasive ventilation in obese patients: a qualitative review and meta-analysis. Surg Obes Relat Dis 12:681–692

13. Muraja-Murro A, Kulkas A, Hiltunen M, Kupari S, Hukkanen T, Tiihonen P, Mervaala E, Töyräs J (2013) The severity of individual obstruction events is related to increased mortality rate in severe obstructive sleep apnea. J Sleep Res 22:663–669

14. Oliveira MG, Treptow EC, Fukuda C, Nery LE, Valadares RM, Tufik S, Bittencourt L, Togeiro SM (2015) Diagnostic accuracy of home-based monitoring system in morbidly obese patients with high risk for sleep apnea. Obes Surg 25:845–851

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

16. Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, Khajehdehi A, Shapiro CM (2008) STOP questionnaire: a tool to screen patients for obstructive sleep apnea. Anesthesiology 108:812–821

17. Marti-Soler H, Hirotsu C, Marques-Vidal P, Vollenweider P, Waeber G, Preisig M, Tafti M, Tufik SB, Bittencourt L, Tufik S, Haba-Rubio J, Heinzer R (2016) The NoSAS score for screening of sleep-disordered breathing: a derivation and validation study. Lancet Respir Med 4:742–748

18. Reed K, Pengo MF, Steier J (2016) Screening for sleep-disordered breathing in a bariatric population. J Thorac Dis 8:268–275

19. Horvath CM, Jossen J, Kröll D, Nett PC, Baty F, Brill AK, Ott SR (2018) Prevalence and prediction of obstructive sleep apnea prior to bariatric surgery-gender-specific performance of four sleep questionnaires. Obes Surg 28:2720–2726

20. Johns MW (1991) A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 14:540–545

21. Chung F, Yang Y, Liao P (2013) Predictive performance of the STOP-Bang score for identifying obstructive sleep apnea in obese patients. Obes Surg 23:2050–2057

22. Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, Marcus CL, Mehra R, Parthasarathy S, Quan SF, Redline S, Strohl KP, Davidson Ward SL, Tangredi MM, American Academy of Sleep Medicine (2012) 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 8:597–619

23. Hanley JA, McNeil BJ (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 148:839–843

24. Duarte RLM, Fonseca LBM, Magalhães-da-Silveira FJ, Silveira EAD, Rabahi MF (2017) Validation of the STOP-Bang questionnaire as a means of screening for obstructive sleep apnea in adults in Brazil. J Bras Pneumol 43:456–463

25. Nagappa M, Liao P, Wong J, Auckley D, Ramachandran SK, Memtsoudis S, Mokhlesi B, Chung F (2015) 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 10:e0143697

26. Glazer SA, Erickson AL, Crosby RD, Kieda J, Zawisza A, Deitel M (2018) The evaluation of screening questionnaires for obstructive sleep apnea to identify high-risk obese patients undergoing bariatric surgery. Obes Surg 28:3544–3552

27. Tan A, Hong Y, Tan LWL, van Dam RM, Cheung YY, Lee CH (2017) Validation of NoSAS score for screening of sleep-disordered breathing in a multiethnic Asian population. Sleep Breath 21:1033–1038

28. Hong C, Chen R, Qing S, Kuang A, Yang H, Su X, Zhao D, Wu K, Zhang N (2018) Validation of the NoSAS score for the screening of sleep-disordered breathing: a hospital-based retrospective study in China. J Clin Sleep Med 14:191–197

29. Vana KD, Silva GE, Goldberg R (2013) Predictive abilities of the STOP-bang and Epworth sleepiness scale in identifying sleep clinic patients at high risk for obstructive sleep apnea. Res Nurs Health 36:84–94

30. Panchasara B, Poots AJ, Davies G (2017) Are the Epworth sleepiness scale and Stop-Bang model effective at predicting the severity of obstructive sleep apnoea (OSA); in particular OSA requiring treatment? Eur Arch Otorhinolaryngol 274:4233–4239