



# An effective model for screening moderate-to-severe obstructive sleep apnea based on the STOP-BANG questionnaire

Yuyu Wang<sup>1,2,3#</sup>, Juanjuan Zou<sup>1,2,4#</sup>, Huajun Xu<sup>1,2</sup>, Cuiping Jiang<sup>5</sup>, Hongliang Yi<sup>1,2</sup>, Jian Guan<sup>1,2</sup>, Shankai Yin<sup>1,2</sup>

<sup>1</sup>Department of Otolaryngology-Head and Neck Surgery, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China; <sup>2</sup>Shanghai Key Laboratory of Sleep Disordered Breathing, Shanghai, China; <sup>3</sup>Department of Otorhinolaryngology-Head and Neck Surgery, Beijing Friendship Hospital, Capital Medical University, Beijing, China; <sup>4</sup>Department of Otorhinolaryngology, Qilu Hospital of Shandong University, NHC Key Laboratory of Otorhinolaryngology (Shandong University), Jinan, China; <sup>5</sup>Department of Endocrinology, Huadong Hospital Affiliated to Fudan University, Shanghai, China

*Contributions:* (I) Conception and design: Y Wang, J Zou, S Yin; (II) Administrative support: H Xu, J Guan; (III) Provision of study materials or patients: Y Wang, C Jiang; (IV) Collection and assembly of data: J Zou, H Yi; (V) Data analysis and interpretation: Y Wang, J Guan; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

*Correspondence to:* Huajun Xu. Department of Otolaryngology-Head and Neck Surgery, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai 200233, China. Email: sunnydayxu2010@163.com; Cuiping Jiang. Department of Endocrinology, Huadong Hospital Affiliated to Fudan University, Shanghai 200040, China. Email: litthat@126.com; Jian Guan. Department of Otolaryngology-Head and Neck Surgery, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai 200233, China. Email: guanjian0606@sina.com.

**Background:** This study aimed to develop a more effective screening model for moderate-to-severe obstructive sleep apnea (OSA) based on the best tool among Epworth Sleepiness Scale (ESS), NoSAS score and STOP-BANG questionnaire (SBQ).

**Methods:** This study screened 2,031 consecutive subjects referred with suspected OSA from 2012 to 2016, including the test cohort from 2012 to 2014 and the validation cohort from 2014 to 2016. Anthropometric measurements, polysomnographic data, ESS, NoSAS scores and SBQ scores were recorded. Receiver operating characteristic curve analyses were performed and the final predictive models were verified in a validation cohort.

**Results:** A total of 1,840 adults were finally included. The performance of ESS, NoSAS score and SBQ in screening OSA was compared. The diagnostic accuracy of SBQ was superior to ESS and NoSAS. A predictive model based on SBQ yielded an area under the curve (AUC) of 0.931 (95% CI: 0.915–0.946), and the sensitivity and specificity were 84.47 (95% CI: 81.4–87.2) and 87.36 (95% CI: 83.9–90.3) respectively. In the validation cohort, the AUC was 0.955 (95% CI: 0.938–0.969), with a sensitivity and specificity of 86.79 (95% CI: 83.2–89.9) and 90.88 (95% CI: 87.2–93.8) respectively. In addition, the model performed moderately in screening mild OSA with the AUC being 0.771 (95% CI: 0.721–0.815).

**Conclusions:** The SBQ was effective in screening moderate-to-severe OSA. And a SBQ -based predictive model afforded excellent diagnostic efficacy, which could be applied in clinical practice.

**Keywords:** Obstructive sleep apnea (OSA); screening model; sensitivity; specificity; STOP-BANG questionnaire (SBQ)

Submitted May 24, 2020. Accepted for publication Jul 01, 2022.

doi: 10.21037/jtd-20-2027

View this article at: <https://dx.doi.org/10.21037/jtd-20-2027>

## Introduction

Obstructive sleep apnea (OSA) is a chronic sleep disorder characterized by a marked reduction or complete cessation of airflow during sleep, affecting 3–24% of the general population (1). It significantly reduces the quality of life and increases the risks of metabolic disorders, hypertension, cardiovascular disease, stroke, and traffic accidents (2–4). Moderate-to-severe OSA, the prevalence of which is 1–14%, is more commonly associated with risks of hypertension, coronary artery disease, and dysfunctional metabolism than mild OSA (5,6), early screening and diagnosis of moderate-to-severe OSA is thus important. However, epidemiological studies have shown that the prevalence of the undiagnosed condition among general populations is as high as 93% in males and 82% in females (7). Although overnight polysomnography (PSG) remains the gold standard for OSA diagnosis, it is difficult to schedule this for snorers given the enormous pressure on hospitalization in China. PSG is expensive and requires a dedicated setup. Moreover, professional staff require extensive training, and the process is time-consuming. All of these factors contribute to the poor diagnosed prevalence. Portable sleep monitor is a cost-effective alternative to PSG, but it is still more expensive and complicated compared with questionnaire survey, and it may underestimate the severity of OSA and potentially yield false-negative results (8). Thus, a simple screening tool for moderate-to-severe OSA is urgently required.

Tools including Epworth Sleepiness Scale (ESS), NoSAS score, and STOP-BANG questionnaire (SBQ) are commonly used to screen for OSA. The ESS yields a subjective measure of excessive daytime sleepiness, and has been widely used to predict OSA (9), but some authors have argued that it is ineffective (10). NoSAS is a new and effective screening tool, but it has been frequently reported to be of insufficient accuracy (11). The SBQ is recognized as one of the best means for OSA screening, but the specificity is unsatisfactory and the results are rather inconsistent (12,13). Also, cutoff applicable to all populations has not been defined (14). New predictive models have recently been reported, but most of them are complicated with poor specificity and inconsistency, which often require biochemical parameters (10,12).

Therefore, we performed a large-scale study to develop a simple and effective screening tool for moderate-to-severe OSA based on the best questionnaire among ESS, NoSAS score and SBQ by evaluating their diagnostic efficiency. We present the following article in accordance with the STARD

reporting checklist (available at <https://jtd.amegroups.com/article/view/10.21037/jtd-20-2027/rc>).

## Methods

### *Study design and population*

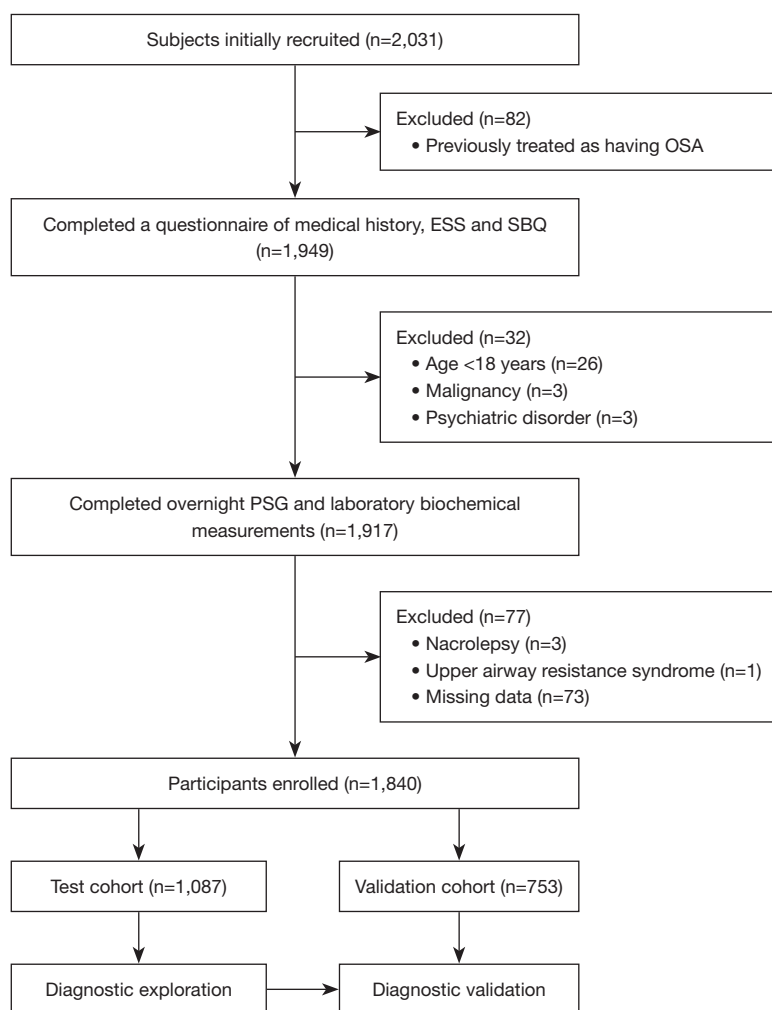
The total cohort consisted of consecutive subjects admitted to the Sleep Center of Shanghai Jiao Tong University Affiliated Sixth People's Hospital with suspected OSA from 2012 to 2016. Consecutive subjects from 2012 to 2014 were enrolled in test cohort and subjects from 2014 to 2016 were enrolled in validation cohort. The inclusion criteria included: (I) subjects were suspected to be suffered from OSA; (II) age  $\geq 18$  years. We excluded those previously treated for OSA, those with malignancy, psychiatric disorder, narcolepsy and upper airway resistance syndrome and those for whom any clinical data were missing (*Figure 1*). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethic Committee of Shanghai Jiao Tong University Affiliated Sixth People's Hospital [No. 2019-KY-050(K)]. Informed consent was obtained from all individual participants included in the study.

### *Anthropometric measurements and physical examination*

Anthropometric measurements including neck circumference (NC), waist circumference (WC) and hip circumference (HC) were measured following standardized protocols. The body mass index (BMI) was the weight divided by the height squared. Blood pressure was measured on the morning after PSG by trained nurses using mercury sphygmomanometers after each patient had rested for at least 5 min in a quiet environment and the result was the average value of three measurements.

### *Questionnaire*

All subjects were requested to complete a questionnaire exploring personal information, alcohol consumption and smoking histories, and any prior illnesses and medical treatments. The ESS explores subjective sleepiness in eight situations (15). Respondents use a four-point scale [0–3] to respond to each of the eight questions, and the scores are summed to give an overall score from 0–24. The NoSAS score has been widely used including Chinese (16,17), and the final score can range from 0–17, with 4 points for



**Figure 1** Flow diagram of recruitment and exclusion of participants. OSA, obstructive sleep apnea; ESS, Epworth Sleepiness Scale; SBQ, STOP-BANG questionnaire; PSG, polysomnography.

an NC  $\geq 40$  cm; 3 for a BMI 25–30 kg/m<sup>2</sup> or 5 for a BMI  $\geq 30$  kg/m<sup>2</sup>; 2 for snoring; 4 for age  $\geq 55$  years; and 2 for being male. The SBQ features eight questions exploring snoring (S), tiredness (T), observed breathing cessation (O), blood pressure (P), BMI (B), age (A), NC, and gender (G) with better performance in Chinese (18,19).

### Overnight PSG

Each subject underwent overnight laboratory-based PSG (Alice 4 or 5 platform; Respironics, Pittsburgh, PA, USA). We performed electroencephalography, electro-oculography, submental and anterior tibialis electromyography, and electrocardiography; we measured

thoracic and abdominal movements, oronasal airflow, and oxygen saturation. All PSG recordings were completed by the same skilled technician using standard criteria. All PSG recordings less than 4 h were repeated or excluded. Respiratory events were carefully scored using the American Academic Sleep Medicine (AASM) criteria. The apnea hypoxia index (AHI) was the number of apnea and hypopnea events per hour of sleep. Subjects with AHI  $\geq 15$  were considered to have moderate-to-severe OSA.

### Statistical analysis

All statistical analysis was performed with the aid of SPSS ver. 21.0 software (SPSS, Chicago, IL, USA). Dichotomous

**Table 1** Basic characteristics of test cohort and validation cohort stratified by gender

Characteristic	Test cohort			Validation cohort		
	Males (n=863)	Females (n=224)	P	Males (n=571)	Females (n=182)	P
Age (years)	41 (33.0–50.0)	48 (39.0–57.0)	<0.001	37 (31.0–49.0)	46 (35.0–56.0)	<0.001
BMI (kg/m <sup>2</sup> )	26.7 (24.5–29.0)	23.9 (22.0–27.3)	<0.001	26.2 (24.2–28.4)	24.2 (21.9–27.6)	<0.001
NC (cm)	40.0 (38.0–42.0)	34.0 (32.6–37.0)	<0.001	40.0 (38.0–42.0)	34.0 (32.9–37.0)	<0.001
WC (cm)	96.0 (90.0–102.0)	86.0 (79.0–94.0)	<0.001	95.0 (90.0–101.0)	86.0 (78.4–95.0)	<0.001
HC (cm)	100.0 (96.0–105.0)	96.0 (91.0–101.0)	<0.001	101.0 (96.5–105.0)	96.0 (91.0–104.0)	<0.001
Sleep data						
AHI (times/h)	29.4 (9.0–55.0)	6.5 (0.8–24.1)	<0.001	27.2 (6.7–55.3)	5.3 (1.2–19.0)	<0.001
Moderate-to-severe OSA (n, %)	567, 65.7	77, 34.4	<0.001	371, 65	53, 29.1	<0.001
LSpO <sub>2</sub> (%)	80.0 (69.0–88.0)	88.5 (80.0–93.0)	<0.001	82.0 (72.0–89.0)	89.0 (82.0–93.1)	<0.001
Mean SpO <sub>2</sub> (%)	94.8 (92.5–96.2)	96.0 (94.0–97.5)	<0.001	90.0 (67.0–95.0)	96.0 (95.0–97.0)	<0.001
Blood data						
FBG (mmol/L)	5.19 (4.86–5.56)	5.04 (4.68–5.54)	0.011	5.18 (4.89–5.60)	5.10 (4.70–5.74)	0.140
TC (mmol/L)	4.67 (4.12–5.34)	4.60 (4.08–5.24)	0.723	4.54 (3.99–5.19)	4.39 (3.70–4.98)	0.015
TG (mmol/L)	1.55 (1.10–2.23)	1.05 (0.75–1.53)	<0.001	1.60 (1.13–2.31)	1.16 (0.75–1.69)	<0.001
HDL (mmol/L)	0.99 (0.87–1.12)	1.19 (1.02–1.40)	<0.001	0.98 (0.87–1.11)	1.12 (0.97–1.30)	<0.001
LDL (mmol/L)	2.90 (2.42–3.42)	2.75 (2.33–3.24)	0.039	2.75 (2.34–3.31)	2.54 (1.97–3.19)	0.001
Questionnaires						
ESS	8.0 (4.0–13.0)	5.0 (2.0–9.8)	<0.001	8.0 (3.0–12.0)	3.0 (0–8.0)	<0.001
NoSAS	7.0 (5.0–11.0)	3.0 (0–6.0)	<0.001	7.0 (4.0–11.0)	3.0 (0–5.0)	<0.001
SBQ	4.0 (3.0–5.0)	1.0 (1.0–3.0)	<0.001	3.0 (2.0–5.0)	1.0 (0–2.0)	<0.001

Dichotomous variables were expressed as percentage, and skewed variables were expressed as medians (interquartile range). The height and weight were recorded in meters and kilograms respectively, with accuracies of 0.01 m and 0.1 kg; NC, WC and HC were recorded in centimeters with accuracies of 0.1 cm. BMI, body mass index; NC, neck circumference; WC, waist circumference; HC, hip circumference; AHI, apnea-hypopnea index; OSA, obstructive sleep apnea; LSpO<sub>2</sub>, lowest oxygen saturation; mean SaO<sub>2</sub>, mean oxygen saturation; FBG, fasting blood glucose; TC, total cholesterol; TG, triglyceride; HDL, high-density lipoprotein; LDL, low-density lipoprotein; ESS, Epworth Sleepiness Scale; SBQ, STOP-BANG questionnaire; NoSAS, neck, obesity, snoring, age, sex.

variables were expressed as percentages and continuous variables as means ± standard deviations (SDs); skewed variables were expressed as medians ± interquartile ranges. The significance of between-group differences in baseline characteristics was examined using the Mann-Whitney U-test, the Wilcoxon rank sum test, or the  $\chi^2$  test. Logistic regression analyses were performed to derive odds ratios with 95% confidence intervals [ORs (95% CIs)] for moderate-to-severe OSA. Receiver operating characteristic (ROC) curves were drawn to assess the diagnostic accuracy of ESS, NoSAS and SBQ evaluations, and that of the predictive model. Sensitivities and specificities were also

recorded. Youden's index (sensitivity + specificity – 1) was used to determine cutoffs maximizing diagnostic utility. A P value <0.05 was considered to be significance.

## Results

### Baseline characteristics

In total, 1,840 consecutive subjects were enrolled, of whom 1,087 were in the test cohort and 753 in the validation cohort (*Figure 1*). The basic characteristics of all subjects are shown in *Table 1*. Males were more obese and tended to

**Table 2** The efficiency of ESS, NoSAS and STOP-BANG in predicting moderate to severe OSAHS

Questionnaire	All (n=1,087)			Males (n=863)			Females (n=224)		
	AUC (95% CI)	Cut-off value	Sensitivity/specificity (%)	AUC (95% CI)	Cut-off value	Sensitivity/specificity (%)	AUC (95% CI)	Cut-off value	Sensitivity/specificity (%)
ESS	0.683 (0.650–0.716)	6	75.2/51.4	0.666 (0.633–0.698)	6	69.4/54.3	0.683 (0.617–0.745)	7	55.4/71.1
NoSAS	0.758 (0.729–0.787)	6.5	68.5/70.2	0.706 (0.669–0.743)	9.5	46.0/84.8	0.805 (0.744–0.867)	4.5	67.5/79.6
SBQ	0.843 (0.820–0.866)	3	88.0/62.8	0.814 (0.787–0.840)	3	72.5/75.7	0.865 (0.813–0.907)	1	85.7/70.8

ESS, Epworth Sleepiness Scale; AUC, area under the curve; OSAHS, obstructive sleep apnea-hypopnea syndrome; CI, confidence interval; SBQ, STOP-BANG questionnaire; NoSAS, neck, obesity, snoring, age, sex.

**Table 3** Performance of the final model in the test and validation cohort

Cohort	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	+LR	-LR
Test cohort	0.931 (0.915–0.946)	84.47 (81.4–87.2)	87.36 (83.9–90.3)	90.7 (88.1–92.9)	79.5 (75.6–83.0)	6.68 (6.4–7.0)	0.18 (0.1–0.2)
Validation cohort	0.955 (0.938–0.969)	86.79 (83.2–89.9)	90.88 (87.2–93.8)	92.5 (89.4–94.9)	84.2 (80.0–87.9)	9.52 (9.0–10.0)	0.15 (0.10–0.2)

AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio.

have a higher AHI and more severe hypoxia than females. Females had lower BMI, NC, WC, and HC, and lower triglyceride and low-density lipoprotein levels than males in both cohorts (all  $P < 0.05$ ). The fasting blood glucose levels significantly differed by gender in test cohort while the difference was not significant in validation cohort. In contrast, the serum level of total cholesterol significantly differed by gender in validation cohort while the difference was not significant in test cohort. The ESS, NoSAS, and SBQ scores were significantly higher in males than females (all  $P < 0.001$ ).

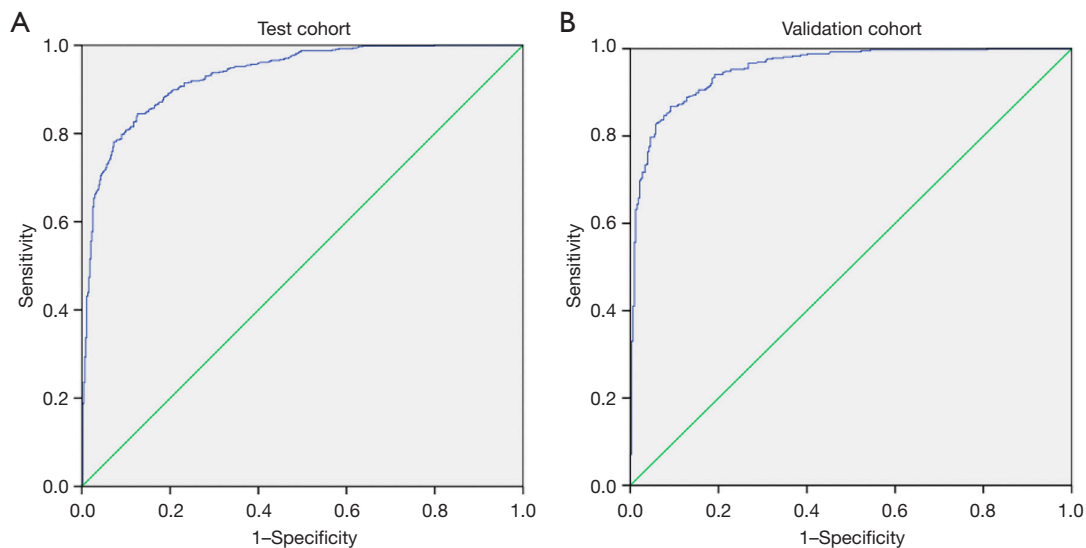
### Performance of ESS, NoSAS, and SBQ

As shown in *Table 2*, the efficacy of SBQ used to predict moderate-to-severe OSA was better than those of the ESS and NoSAS; the area under the curve (AUC) was 0.843 (95% CI: 0.820–0.866), with a sensitivity 88.0% and a specificity 62.8%. After stratification by gender, SBQ remained optimal, with AUCs of 0.814 (95% CI: 0.787–0.840) in males and 0.865 (95% CI: 0.813–0.907) in females. A gender-related difference in optimal diagnostic SBQ cutoffs was significant, the values of which were 3 for males and 1 for females (*Table 2*).

### A predictive model based on SBQ

As SBQ yielded the optimal diagnostic accuracy in terms of predicting moderate-to-severe OSA, we developed a predictive model to maximize diagnostic accuracy. Due to the fact that BMI, age, gender, and NC were all included in SBQ, the model incorporated other parameters including WC, HC, the lowest  $\text{SpO}_2\%$  ( $\text{LSpO}_2$ ) other than SBQ score. After multiple regression analysis, the predictive model included WC,  $\text{LSpO}_2$ , and the SBQ score; the Logit (P) was  $0.449 \times \text{SBQ} + 0.023 \times \text{WC} - 0.194 \times \text{LSpO}_2 + 12.787$ . The AUC of the constructed model was 0.931 (95% CI: 0.915–0.946), and the sensitivity and specificity were 84.47 (95% CI: 81.4–87.2) and 87.36 (95% CI: 83.9–90.3) respectively (*Table 3* and *Figure 2A*). Then the model was testified in the validation cohort. The AUC was 0.955 (95% CI: 0.938–0.969), with a sensitivity and specificity of 86.79 (95% CI: 83.2–89.9) and 90.88 (95% CI: 87.2–93.8) respectively (*Table 3* and *Figure 2B*).

Considering the possibility of mild OSA aggravation, screening mild patients is also important. Therefore, we investigated the performance of the model in screening overall OSA including mild patients. As shown in *Table S1*, the model performed excellently in screening



**Figure 2** ROC curves for the final model in the test cohort and validation cohort. (A) The AUC in the test cohort was 0.931 (95% CI: 0.915–0.946); (B) The AUC in the test cohort was 0.955 (95% CI: 0.938–0.969). ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval.

overall OSA with the AUC being 0.926 (95% CI: 0.904–0.943), and performed moderately in screening mild OSA with the AUC being 0.771 (95% CI: 0.721–0.815).

## Discussion

In the present large-scale study, we found that the SBQ was effective when used to screen for moderate-to-severe OSA, and it performed better than the ESS and NoSAS in this context. The predictive model based on SBQ, WC and  $LpO_2$  afforded excellent diagnostic efficiencies.

Given the severity of the condition and the low diagnosis rate, some studies have focused on screening for moderate-to-severe OSA using questionnaires such as the ESS, NoSAS, and SBQ (20,21). ESS is often used to assess subjective sleepiness (15); higher scores are associated with an increased risk of metabolic syndrome in males with OSA (22). Although the ESS is widely used to screen for OSA, its diagnostic reliability has been questioned. Kum *et al.* (23) found that ESS scores increased as AHI thresholds increased, but some studies have found weak or no correlations between ESS scores and OSA severity, suggesting that the ESS may not aid in OSA diagnosis (10,24). We found that the ESS did not reliably predict moderate-to-severe OSA, in agreement with Miller and Ghandeharioun *et al.* (10,24). Maybe it was because that the ESS is totally subjective and influenced by many factors

including daytime naps and blood pressure (25,26).

The NoSAS, which measures five items objectively, is a new and effective screening tool affording better performance, and has been validated in an external population (16,17). It was demonstrated that the NoSAS exhibited good predictive performances in both European and Asian populations (17,27). However, Giampá *et al.* (11) found that the NoSAS was inaccurate when used to diagnose OSA in patients with resistant hypertension. We found that the NoSAS yielded a better AUC but had a low sensitivity in terms of predicting moderate and severe OSA, perhaps because some subjective symptoms are also usefully predictive.

SBQ is a part-self-administered questionnaire based on the Berlin instrument, and includes both subjective symptoms and objective indicators; the design is thus more reasonable than that of the ESS. SBQ is an acronym derived from the first letters of the questions (18,19), facilitating completion of the instrument. Many researchers consider SBQ to be one of the best measures for OSA screening. Chiu *et al.* (20), in a meta-analysis of various questionnaires, concluded that SBQ optimally screened for OSA. A recent report also indicated that SBQ was diagnostically more accurate than the ESS (10). However, similar to most questionnaires, SBQ affords high sensitivity but only low specificity when diagnosing mild, moderate, and severe OSA (12,13), and no one cutoff is appropriate for all

populations (14,18). We found that SBQ was moderately effective when used to predict moderate-to-severe OSA, and the performance was better than ESS and NoSAS.

An AASM task force evaluated the current screening tools, and concluded that no single tool was ideal (28). Therefore, development of a predictive model affording adequate sensitivity and specificity became urgent. Simpson *et al.* (29) reported that a combination of snoring frequency and hypertension status could be used to identify those with moderate-to-severe OSA in general populations, but the sample size was small and diagnosis of moderate-to-severe OSA was not PSG-based. Sun *et al.* (30) used an artificial intelligence model to screen for moderate-to-severe OSA. The sensitivity and specificity were satisfactory, but only 120 patients were evaluated. We built a new model based on SBQ data, WC and LSpO<sub>2</sub> that are readily available. The AUCs, sensitivities, and specificities for both males and females were good. In the validation cohort, model performance was excellent.

Although we enrolled a relatively large number of subjects, our work had certain limitations. Firstly, the prevalence of OSA in subjects referred to our sleep center was higher than that in the general population, which may compromise the generalization of our model. Secondly, ethnicity or menopausal status may affect sleepiness in females with suspected OSA. Significantly, in our sample, the percentage and severity of female OSA were much lower than that of male OSA, but the situation may be different in other samples and the predictive value will be affected as the prevalence of moderate to severe OSA changes among samples. Therefore, the diagnostic efficiency of our model must be further verified in other samples including different country of residence and ethnicity.

In conclusion, the SBQ performed excellently in screening moderate-to-severe OSA. And a predictive model that combined SBQ with WC and LpO<sub>2</sub> afforded excellent diagnostic efficiency, which provided a simple and effective screening tool for OSAHS.

## Acknowledgments

We thank all of the participants and survey staffs for their participation. And we thank the language editing service by Textcheck.

**Funding:** This study was funded by National Key R&D Program of China (No. 2017YFC0112500), National Natural Science Foundation of China (Nos. 81770987, 81700896, 81701306, and 81770988), Shanghai

Municipal Commission of Science and Technology (No. 18DZ2260200), Shanghai Shenkang Hospital Development Center (No. 16CR3103B), Innovation Program of Shanghai Municipal Education Commission (No. 2017-01-07-00-02-E00047), multi-center clinical research project from school of medicine, Shanghai Jiao Tong University (No. DLY201502) and Shanghai Shen-Kang Hospital Management Center Project (No. SHDC12015101).

## Footnote

**Provenance and Peer Review:** This article was a standard submission to the series “Sleep Section” published in the *Journal of Thoracic Disease*. The article has undergone external peer review.

**Reporting Checklist:** The authors have completed the STARD reporting checklist. Available at <https://jtd.amegroups.com/article/view/10.21037/jtd-20-2027/rc>

**Data Sharing Statement:** Available at <https://jtd.amegroups.com/article/view/10.21037/jtd-20-2027/dss>

**Peer Review File:** Available at <https://jtd.amegroups.com/article/view/10.21037/jtd-20-2027/prf>

**Conflicts of Interest:** All authors have completed the ICMJE uniform disclosure form (available at <https://jtd.amegroups.com/article/view/10.21037/jtd-20-2027/coif>). The series “Sleep Section” was commissioned by the editorial office without any funding or sponsorship. The authors have no other conflicts of interest to declare.

**Ethical Statement:** The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethic Committee of Shanghai Jiao Tong University Affiliated Sixth People’s Hospital [No. 2019-KY-050(K)]. Informed consent was obtained from all individual participants included in the study.

**Open Access Statement:** This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-

commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

- Berg S. Obstructive sleep apnoea syndrome: current status. *Clin Respir J* 2008;2:197-201.
- Bergeron M, Ishman SL. Persistent Obstructive Sleep Apnea Burden on Family Finances and Quality of Life. *Otolaryngol Head Neck Surg* 2021;165:483-9.
- Salman LA, Shulman R, Cohen JB. Obstructive Sleep Apnea, Hypertension, and Cardiovascular Risk: Epidemiology, Pathophysiology, and Management. *Curr Cardiol Rep* 2020;22:6.
- Zou J, Xia Y, Xu H, et al. Independent relationships between cardinal features of obstructive sleep apnea and glycometabolism: a cross-sectional study. *Metabolism* 2018;85:340-7.
- Hirotsu C, Haba-Rubio J, Togeiro SM, et al. Obstructive sleep apnoea as a risk factor for incident metabolic syndrome: a joined Episono and HypnoLaus prospective cohorts study. *Eur Respir J* 2018;52:1801150.
- Ishiwata S, Tomita Y, Ishiwata S, et al. Association between Obstructive Sleep Apnea and SYNTAX Score. *J Clin Med* 2020;9:3314.
- Ancoli-Israel S, Kripke DF, Klauber MR, et al. Sleep-disordered breathing in community-dwelling elderly. *Sleep* 1991;14:486-95.
- Goldstein CA, Karnib H, Williams K, et al. The utility of home sleep apnea tests in patients with low versus high pre-test probability for moderate to severe OSA. *Sleep Breath* 2018;22:641-51.
- Senaratna CV, Perret JL, Lowe A, et al. Detecting sleep apnoea syndrome in primary care with screening questionnaires and the Epworth sleepiness scale. *Med J Aust* 2019;211:65-70.
- 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.
- Giampá SQC, Pedrosa RP, Gonzaga CC, et al. Performance of NoSAS score versus Berlin questionnaire for screening obstructive sleep apnoea in patients with resistant hypertension. *J Hum Hypertens* 2018;32:518-23.
- Reuter H, Herkenrath S, Treml M, et al. Sleep-disordered breathing in patients with cardiovascular diseases cannot be detected by ESS, STOP-BANG, and Berlin questionnaires. *Clin Res Cardiol* 2018;107:1071-8.
- Abumuamar AM, Dorian P, Newman D, et al. The STOP-BANG questionnaire shows an insufficient specificity for detecting obstructive sleep apnea in patients with atrial fibrillation. *J Sleep Res* 2018;27:e12702.
- Chung F, Yang Y, Brown R, et al. Alternative scoring models of STOP-bang questionnaire improve specificity to detect undiagnosed obstructive sleep apnea. *J Clin Sleep Med* 2014;10:951-8.
- Peng LL, Li JR, Sun JJ, et al. Reliability and validity of the simplified Chinese version of Epworth sleepiness scale. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* 2011;46:44-9.
- Qing SM, Chen RK, Liu H, et al. Comparison of the NoSAS score with four different questionnaires as screening tools for obstructive sleep apnea-hypopnea syndrome. *Zhonghua Jie He He Hu Xi Za Zhi* 2018;41:213-9.
- 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;21:1033-8.
- Wang W, Yuan S, Le Grange JM, et al. Evaluating the performance of five scoring systems for prescreening obstructive sleep apnea-hypopnea syndrome. *Sleep Breath* 2021;25:1685-92.
- Waseem R, Chan MTV, Wang CY, et al. Diagnostic performance of the STOP-Bang questionnaire as a screening tool for obstructive sleep apnea in different ethnic groups. *J Clin Sleep Med* 2021;17:521-32.
- Chiu HY, Chen PY, Chuang LP, et al. Diagnostic accuracy of the Berlin questionnaire, STOP-BANG, STOP, and Epworth sleepiness scale in detecting obstructive sleep apnea: A bivariate meta-analysis. *Sleep Med Rev* 2017;36:57-70.
- Arora T, Al-Houqani M. Comparison of commonly used screening tools for determining obstructive sleep apnea amongst aviation employees. *Sleep Med* 2021;77:332-6.
- Fu Y, Xu H, Xia Y, et al. Excessive daytime sleepiness and metabolic syndrome in men with obstructive sleep apnea: a large cross-sectional study. *Oncotarget* 2017;8:79693-702.
- Kum RO, Özcan M, Yurtsever Kum N, et al. A new suggestion for the Epworth Sleepiness Scale in obstructive sleep apnea. *Eur Arch Otorhinolaryngol* 2015;272:247-52.
- Ghandeharioun H, Rezaeitab F, Lotfi R. Analysis of respiratory events in obstructive sleep apnea syndrome:



- Inter-relations and association to simple nocturnal features. *Rev Port Pneumol* (2006) 2016;22:86-92.
25. Saletin JM, Hilditch CJ, Dement WC, et al. Short Daytime Naps Briefly Attenuate Objectively Measured Sleepiness Under Chronic Sleep Restriction. *Sleep* 2017;40:zsx118.
  26. Dutheil F, Danini B, Bagheri R, et al. Effects of a Short Daytime Nap on the Cognitive Performance: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health* 2021;18:10212.
  27. Horvath CM, Jossen J, Kröll D, et al. Prevalence and Prediction of Obstructive Sleep Apnea Prior to Bariatric Surgery-Gender-Specific Performance of Four Sleep Questionnaires. *Obes Surg* 2018;28:2720-6.
  28. Gamaldo C, Buenaver L, Chernyshev O, et al. Evaluation of Clinical Tools to Screen and Assess for Obstructive Sleep Apnea. *J Clin Sleep Med* 2018;14:1239-44.
  29. Simpson L, Hillman DR, Cooper MN, et al. High prevalence of undiagnosed obstructive sleep apnoea in the general population and methods for screening for representative controls. *Sleep Breath* 2013;17:967-73.
  30. Sun LM, Chiu HW, Chuang CY, et al. A prediction model based on an artificial intelligence system for moderate to severe obstructive sleep apnea. *Sleep Breath* 2011;15:317-23.

**Cite this article as:** Wang Y, Zou J, Xu H, Jiang C, Yi H, Guan J, Yin S. An effective model for screening moderate-to-severe obstructive sleep apnea based on the STOP-BANG questionnaire. *J Thorac Dis* 2022;14(8):3066-3074. doi: 10.21037/jtd-20-2027

## Supplementary

**Table S1** Performance of the final model

Variable	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	+LR	-LR
OSA	0.926 (0.904–0.943)	82.59 (79.1–85.7)	87.32 (82.1–91.5)	94.3 (91.8–96.2)	66.4 (60.6–71.9)	6.52 (6.1–6.9)	0.20 (0.1–0.3)
Mild OSA	0.771 (0.721–0.815)	84.48 (76.6–90.5)	62.44 (55.6–69.0)	55.1 (47.4–62.5)	88.1 (81.8–92.8)	2.25 (2.0–2.6)	0.25 (0.2–0.4)

AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; +LR, positive likelihood ratio; -LR, negative likelihood ratio.