Nature and Science of Sleep

Elevate Series

Open Access Full Text Article

ORIGINAL RESEARCH

Performance of Four Screening Tools for Identifying Obstructive Sleep Apnea Among Patients with Insomnia

Chuan Shi*, Yuxin Wang*, Jinmei Luo, Rong Huang, Yi Xiao

Department of Pulmonary and Critical Care Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Science (CAMS) and Peking Union Medical College (PUMC), Beijing, 100730, People's Republic of China

*These authors contributed equally to this work

Correspondence: Jinmei Luo, Department of Pulmonary and Critical Care Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Science (CAMS) and Peking Union Medical College (PUMC), Beijing, 100730, People's Republic of China, Email palmljm@126.com

Purpose: Co-morbid insomnia and sleep apnea are common in clinical practice. The existing OSA screening tools have not been fully validated in insomnia populations, and items measuring daytime function may be interfered with the presence of insomnia. This study aims to validate the performance of four commonly used OSA screening tools among individuals with and without insomnia.

Participants and Methods: A cross-sectional survey was conducted in individuals with suspected OSA referred for sleep studies from December 2021 to December 2023. All participants completed the Insomnia Severity Index (ISI) scale, STOP-Bang, Epworth Sleepiness Scale (ESS), Berlin questionnaire, and NoSAS score. Clinical insomnia was defined as an ISI of 15 or more. Performance of screening tools was primarily assessed by sensitivity, specificity, and the receiver operating characteristic (ROC) curve.

Results: A total of 1266 participants (26% females, age 46.4 ± 12.4 years) were included in the study. The prevalence of apnea-hypopnea index (AHI) ≥15/h was 48% and 52% in the insomnia (n=313) and non-insomnia (n=953) group, respectively (P>0.05). In presence of insomnia, the STOP-Bang, ESS, and Berlin questionnaire demonstrated higher sensitivity but lower specificity. Using conventional cutoffs, the STOP-Bang had the highest level of sensitivity (93.2%, 95% CI 87.6–96.5%), while NoSAS had the highest level of specificity (67.7%, 95% CI 59.9–74.6%) for identifying AHI ≥15/h. The STOP-Bang and NoSAS outperformed ESS and Berlin with areas under the ROC curve >0.7 at all levels of OSA severity. The Youden's index was maximized at score 4 for STOP-Bang and score 7 for NoSAS.

Conclusion: The performance of OSA screening tools incorporating evaluation of daytime function is altered in the presence of insomnia. Under conventional cutoffs, STOP-Bang is the preferred screening tool due to its high sensitivity.

Keywords: obstructive sleep apnea, insomnia, screening instruments, polysomnography

Introduction

Obstructive sleep apnea (OSA) is a clinical condition caused by repeated episodes of upper airway collapses during sleep. The resultant sleep disruption is frequently associated with excessive daytime sleepiness. Untreated OSA may increase the risk of cardiovascular disorders, stroke, diabetes mellitus, cognitive impairment, traffic accidents, and ultimately damage the quality of life and life expectancy.[1] It has been estimated that over 900 million individuals are affected by OSA globally, creating a massive socioeconomic burden.[2]

To date, polysomnography (PSG) and home sleep apnea test (HSAT) are the standard medical test for the diagnosis of OSA.[3] However, the time-consuming and technology-dependent PSG test often hinders the assessment.[4] While the HSAT provides a more feasible technique, the resources for sleep services are still inadequate in vast regions where health resource is limited, leaving a large proportion of OSA cases unrecognized.[5] To cope with this situation, numerous screening tools have been developed.[6] These tools rapidly identify individuals at high risk for OSA and guide subsequent diagnostic studies in general clinical settings. The adaptability of these tools for specific at-risk OSA populations demands validation.

Received: 5 September 2024
Accepted: 17 February 2025
Published: 3 March 2025

Co-morbid insomnia and sleep apnea (COMISA) has attracted much attention as it has been linked to higher levels of psychological distress, greater cardiovascular and metabolic risk, and increased all-cause mortality.[7–9] The pooled prevalence of insomnia among patients with OSA was 38%, and up to 35% of insomnia patients fulfilled the diagnostic criteria for OSA.[10] This higher-than-expected prevalence of the comorbidity, together with other experimental and clinical evidence, has suggested an interplay between the two conditions, leading to a cycle of poor sleep quality and impaired daytime function.[7]

The scale of excessive daytime sleepiness has been incorporated in some of the most frequently used OSA screening questionnaires, including STOP-Bang,[11] Berlin,[12] and Epworth sleepiness scale (ESS).[13] However, nearly half of insomnia sufferers may also experience excessive daytime sleepiness, questioning the performance of such tools in this population.[14] To date, experience of OSA screening among individuals with insomnia remains scarce.[15] The current study was designed to evaluate the performance of STOP-Bang, Berlin, ESS, as well as NoSAS (a newly proposed screening tool without items associated with daytime sleepiness) in a sample of consecutive adult patients with or without insomnia.

Methods

Participants

This cross-sectional study was conducted at a sleep breathing disorder clinic in northern China.

Individuals were included if they had been suspected of OSA and referred to in-laboratory PSG. Individuals were excluded for any of the following reasons: 1) age <18 years, 2) received treatment for sleep apnea prior to enrollment, 3) unable to fill out or did not complete the electronic questionnaire; 5) technically inadequate PSG. All participants provided written consent. The study was approved by the Research Ethics Committee of Peking Union Medical College Hospital (Approval No. K24C2001) and performed in accordance with the Declaration of Helsinki.

As the study was designed to compare primarily the sensitivity of screening tools, the sample size required for participants with insomnia was calculated through the equation $n = \frac{\left[Z_\alpha\sqrt{\Psi}+Z_\beta\sqrt{\Psi-(Se_2-Se_1)^2}\right]^2}{(Se_2-Se_1)^2}$ where n is the sample size, Se_1 and Se_2 are the reported sensitivity of two close performing tools (STOP-Bang and NoSAS) for identifying moderate-severe OSA, and Ψ is the probability of disagreement between the two tools.[15,16] Here we applied anα of 5%, β of 10%, and maximum Ψ of $Se_1 \times (1 - Se_2) + Se_2 \times (1 - Se_1)$. Assuming one-fourth of patients with suspected OSA has comorbid insomnia, the final enrollment size was estimated to be at least 828. Considering the annual workload of our center, we enrolled participants from January 2022 to December 2023.

A total of 1428 individuals with suspected OSA were referred for in-lab PSG during the study period, among which 162 individuals (11%) were excluded according to the criteria. The number of individuals excluded due to each exclusion criteria was shown in Figure 1.

Before performing PSG, patient characteristics including age, sex, marital status, smoking and drinking status, and use of sleep aid were collected. Information on underlying clinical conditions was collected both from patient's self-reports as well as the electronic medical record system. On the night of PSG, an integrated electronic form that consisted: 1) Insomnia Severity Index (ISI); 2) STOP-Bang questionnaire; 3) Berlin questionnaire; 4) Epworth Sleepiness Scale (ESS); 5) NoSAS score was then filled out by each participant. Assistance, if needed, was provided by a sleep physician during the survey. The height, body mass, and neck circumference (NC) were then measured. The body-mass index (BMI) was calculated by dividing the weight by the square of the height.

Insomnia Assessment

The ISI is a validated and globally employed questionnaire assessing the nature, severity, and impact of insomnia.[17] The self-reported instrument is composed of 7 items that evaluate 1) the severity of sleep-onset, 2) sleep maintenance, 3) early morning awakening, 4) satisfaction with current sleep pattern, 5) interference with daily functioning, 6) noticeability of impairment attributed to the sleep problem, 7) level of distress caused by the sleep problem during the last two weeks. Each of these items is rated on a 5-point Likert scale (eg, 0 = no problem, 4 = very severe problem), yielding

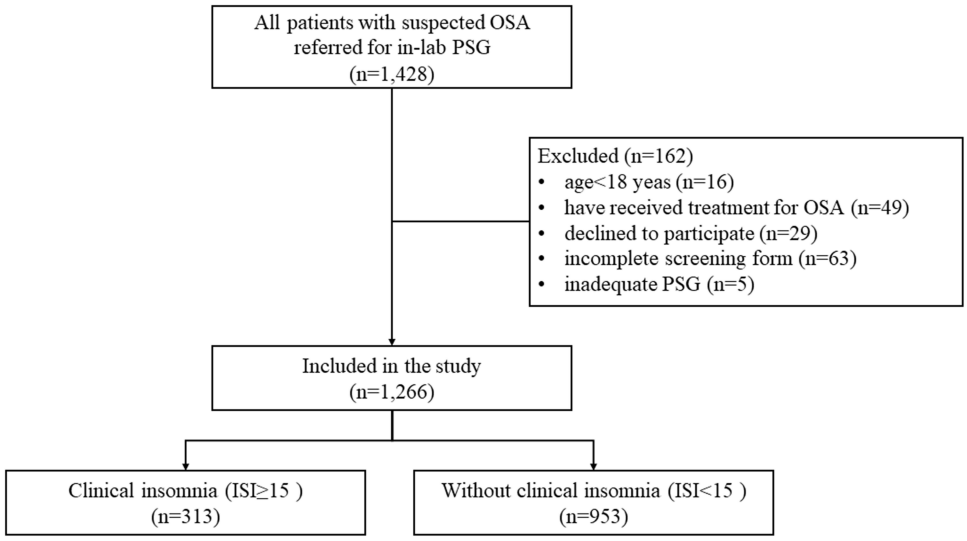

Figure 1 STROBE flowchart of the study.
Abbreviation: ISI, Insomnia Severity Index.

a total score from 0 to 28. A total score ≥15 had been proposed by the developers as "clinically significant insomnia".[17] This cutoff was adopted by the current study for the grouping of participants. For simplicity, the group with ISI ≥15 was referred to as "clinical insomnia", while the group with ISI <15 was referred to as "non-insomnia" in the text. This only reflects the severity and impact of insomnia and does not imply a clinical diagnosis.

OSA Screening

The STOP-Bang questionnaire includes 8 yes or no items: "S" for loud snoring, "T" for tiredness during daytime, "O" for observed apnea, "P" for high blood pressure, "B" for BMI >35 kg/m^2, "a" for age >50 years, "n" for NC > 40 cm, and "g" for male gender.[11] One point is given for each question answered with yes, and a total score of 3 or more sensitively predicts OSA in both the original study and a validation study among the Chinese population.[11,18]

The Berlin questionnaire consists of 10 questions with additional information on height and weight. These items were arranged into 3 categories: 1) snoring and cessation of breathing (5 questions); 2) excessive daytime sleepiness (3 questions); 3) BMI and hypertension (1 question plus BMI). Positive scores in 2 or more categories suggest a high risk for OSA.[12]

The ESS measures the level of daytime sleepiness by the chances of falling asleep in 8 daily conditions. Each item is rated from 0 (would never doze) to 3 (high chance of dozing), totaling a score from 0 to 24.[13] ESS score above 10 has been widely recognized as excessive daytime sleepiness which motivates further OSA evaluation in patients with snoring.[19]

The NoSAS is a clinical score that allocates 4 points for NC > 40 cm, 3 points for 25 kg/m^2 < BMI < 30 kg/m^2 or 5 points for BMI > 30 kg/m^2, 2 points for snoring, 4 points for age > 55 years, and 2 points for being male.[20] A threshold of 8 points was chosen in the original study to identify individuals at risk for clinically significant OSA. The utility of NoSAS has been validated in a sleep clinic-based study in China.[21]

Among the above tools, daytime function is weighted as follows: 8/8 items in ESS, 1/3 categories in Berlin questionnaire, 1/8 items in STOP-Bang questionnaire, and none of the items in NoSAS.

Polysomnography

All participants underwent full-night PSG (Embla N7000; Natus Medical Incorporated, Orlando, USA; or SOMNO HD; SOMNO Medics AG, Randersacker, Germany) from 10 PM to 6 AM. Electroencephalogram, electrooculogram, chin electromyogram, electrocardiogram, airflow, respiratory effort, pulse oxygen saturation (SpO_2), snoring, and body position were continuously recorded. Data were recorded on software that accompanies the device (RemLogic for Embla N7000; DOMINO for SOMNO HD) and scored by two experienced sleep technicians following the AASM Manual for the Scoring of Sleep and Associated Events (version 2.6). The technicians were blinded for insomnia status as

https://doi.org/10.2147/NSS.S494804

well as OSA screening results. The documented Cohen's kappa in sleep staging between the two technicians was 0.80 [95% CI 0.75, 0.84].

Apnea was scored as a drop of airflow ≥90% from the pre-event baseline lasting at least 10 seconds. Hypopnea was defined as a drop of airflow ≥30% lasting at least 10 seconds with an associated ≥3% decrease in SpO_2 or arousal response of more than 3 seconds. AHI was defined as the number of apnea and hypopnea events per hour of sleep. Oxygen desaturation index (ODI) was defined as the number of SpO_2 drops of 3% or more per hour of sleep. Total sleep time (TST), sleep latency, sleep efficiency, wake after sleep onset (WASO), the percentage of each sleep stage, the lowest oxygen saturation by pulse oximetry ($LSpO_2$), and the percentage of sleep time with SpO_2 below 90% (T90) were also collected.

Data Analysis

Data were analyzed with SPSS v. 22.0 software (IBM, Armonk, NY, United States). The normality of variables was tested by the Kolmogorov–Smirnov test. Demographic, clinical, and PSG data were summarized either as means and standard deviation (quantitative data in normal distribution), median and interquartile (quantitative data not in normal distribution), or exact counts and percentages (qualitative data). Comparisons between the insomnia and non-insomnia groups were made using Student's *t*-test or Mann–Whitney *U*-test for continuous variables, and chi-square test for categorical variables. Correlations between scores of the screening tools and AHI were examined by Pearson's (STOP-Bang, ESS, and NoSAS) or Spearman's (Berlin) correlation analysis. Using PSG data as the gold standard, the sensitivities, specificities, positive predictive value (PPV), and negative predictive value (NPV) of each screening tool were calculated. The diagnostic performance of screening tools was further assessed through the net reclassification improvement (NRI), area under the receiver operating characteristic (ROC) curves and Youden's index (YI). The NRI and its significance testing were calculated by the method proposed by Pencina et al.[22] The areas under the ROC curves (AUCs) were compared with the DeLong method using the MedCalc software (MedCalc Software Ltd, Ostend, Belgium).[23] All statistical tests were 2-sided unless specified, and *P* values less than 0.05 were considered statistically significant.

Results

Study Population

Of the 1266 eligible patients, 313 were classified as clinical insomnia. Characteristics of the insomnia and non-insomnia groups are shown in Table 1. Demographic and clinical characteristics were comparable between the two groups except for a higher proportion of females and use of sleep aids among participants with insomnia.

No significant difference was identified regarding the means of AHI and ODI between the two groups. The prevalence of AHI≥5/h, AHI≥15/h, AHI≥30/h was 72%, 48%, 28% for the insomnia group, and 76%, 52%, 34% for the non-insomnia group, respectively. There was no significant difference in the frequency between insomnia- and non-insomnia patients ($P>0.05$ in all comparisons).

Screening Scores in Different Insomnia Severity Groups

The STOP-Bang and NoSAS scores did not differ significantly between the insomnia and non-insomnia groups (Table 1). The ESS and number of positive categories in the Berlin questionnaire were greater in the clinical insomnia group.

The contribution of items related to daytime function in the STOP-Bang and Berlin questionnaire was then investigated. The positive rate for "tired" in the STOP-Bang was 52% in the insomnia group and 49% in the non-insomnia group ($P>0.05$). Notably, the positive rate for category 2 (excessive daytime sleepiness) in Berlin questionnaire was 50% in the insomnia group, compared to only 17% in the non-insomnia group ($P<0.0001$). The total score and positivity of category 1 (snoring and cessation of breathing) and category 3 (BMI and hypertension) of Berlin questionnaire were comparable between the insomnia and non-insomnia groups ($P>0.05$ in all comparisons).

Using conventional cutoffs, high risk for OSA was present in 81% of insomnia and 75% of non-insomnia participants defined by STOP-Bang ($P=0.02$), 65% of insomnia and 46% of non-insomnia participants defined by ESS ($P<0.0001$), 63% of insomnia and 43% of non-insomnia participants defined by Berlin questionnaire ($P<0.0001$), and 48% of insomnia and 52% of non-insomnia participants defined by NoSAS ($P>0.05$) (Table 1).

Table 1 Description of the Study Population

	Non-insomnia (ISI<15), n=953	Clinical Insomnia (ISI≥15), n=313	*P*
Demographic and clinical characteristics			
Age (years), mean±SD	45.8±12.3	47.8±12.6	0.89
Female, n (%)	232 (24)	102 (33)	0.004
Marital status, n (%)			0.09
Single	125 (13)	42 (13)	
Married	792 (83)	250 (80)	
Divorced/separated/widowed	36 (4)	21 (7)	
Current smoker, n (%)	183 (19)	57 (18)	0.698
Regular use of a sleep aid, n (%)	27 (3)	45 (14)	<0.001
BMI (kg/m^2), mean±SD	26.8±4.5	26.3±4.6	0.81
Neck circumference (cm), mean±SD	39.3±4.0	38.9±4.3	0.072
Hypertension, n (%)	299 (31)	110 (35)	0.216
Diabetes mellitus, n (%)	95 (10)	42 (13)	0.088
Cardiovascular disease, n (%)	143 (15)	59 (19)	0.107
OSA-related questionnaire			
STOP-Bang (points), mean±SD	3.5±1.5	3.9±1.5	0.566
STOP-Bang ≥ 3 points, n (%)	711 (75)	252 (81)	0.02
ESS (points), mean±SD	9.4±5.5	12.5±6.3	0.001
ESS ≥ 11 points, n (%)	406 (43)	197 (63)	<0.001
Berlin (positive categories), median [IQR]	1 [1, 2]	2 [1, 2]	<0.001
Berlin ≥ 2 positive categories, n (%)	438 (46)	202 (65)	<0.001
NoSAS (points), mean±SD	8.1±4.0	7.6±4.0	0.971
NoSAS ≥ 8 points, n (%)	492 (52)	150 (48)	0.256
Polysomnographic data			
Total sleep time (min), mean±SD	384.0±63.0	371.8±65.7	0.162
Sleep efficiency (%), mean±SD	87.7±12.0	86.3±13.2	0.016
Awakenings (n), median [IQR]	6 [4, 10]	6 [3, 9]	0.212
WASO (min), median [IQR]	38.2 [14.5, 76.6]	41.6 [16.0, 93.2]	0.149
Sleep latency (min), median [IQR]	15.0 [6.5, 34.5]	17.5 [7.3, 39.7]	0.126
Stage N1 sleep (%), mean±SD	9.1±6.9	8.3±5.9	0.06
Stage N2 sleep (%), mean±SD	43.2±16.8	42.0±14.8	0.738
Stage N3 sleep (%), mean±SD	21.7±13.6	22.4±13.8	0.575
Stage R sleep (%), mean±SD	14.3±6.2	13.7±6.1	0.268
AHI (n/h), mean±SD	24.9±24.5	22.4±24.4	0.293
ODI (n/h), mean±SD	21.6±23.3	19.6±23.5	0.318
$LSpO_2$ (%), mean±SD	85.6±8.7	86.0±8.9	0.442
T90 (%), median [IQR]	0 [0, 1.2]	0 [0, 1.1]	0.288

Abbreviations: AHI, apnea/hypopnea index; BMI, body mass index; ESS, Epworth sleepiness scale; IQR, interquartile range; ISI, insomnia severity index; $LSpO_2$, lowest pulse oxygen saturation; ODI, oxygen desaturation index; REM, rapid eye movement; T90, time under 90% oxygen saturation; WASO, wake after sleep onset.

Correlation Between Screening Scores and Polysomnographic Data

The correlation between screening scores (positive categories in the case of the Berlin questionnaire) and PSG parameters that reflect the severity of OSA (ie, AHI, ODI, $LSpO_2$, T90) was evaluated among participants with and without insomnia (Table 2). The STOP-Bang and NoSAS scores demonstrated low but significant correlation with AHI, ODI, and $LSpO_2$ with and without the presence of clinical insomnia. The correlation between ESS or Berlin questionnaire with PSG data was relatively weak, as in both groups *r* or Spearman rho were <0.3 or around 0.3 for all variables.

https://doi.org/10.2147/NSS.S494804

Table 2 Correlation Analysis Between Screening Scores and Polysomnographic Data

	STOP-Bang		ESS		Berlin		NOSAS	
	r	*P*	*r*	*P*	Spearman *ρ*	*P*	*r*	*P*
Non-insomnia (ISI<15), n=953								
AHI	0.401	<0.001	0.171	<0.001	0.291	<0.001	0.396	<0.001
ODI	0.402	<0.001	0.160	<0.001	0.309	<0.001	0.415	<0.001
$LSpO_2$	−0.374	<0.001	−0.158	<0.001	−0.266	<0.001	−0.350	<0.001
T90	0.218	<0.001	0.135	<0.001	0.220	<0.001	0.194	<0.001
Clinical insomnia (ISI≥15), n=313								
AHI	0.427	<0.001	0.261	<0.001	0.240	<0.001	0.446	<0.001
ODI	0.409	<0.001	0.247	<0.001	0.267	<0.001	0.453	<0.001
$LSpO_2$	−0.409	<0.001	−0.283	<0.001	−0.228	<0.001	−0.444	<0.001
T90	0.267	<0.001	0.225	<0.001	0.231	<0.001	0.238	<0.001

Abbreviations: AHI, apnea/hypopnea index; ESS, Epworth sleepiness scale; ISI, insomnia severity index; $LSpO_2$, lowest pulse oxygen saturation; ODI, oxygen desaturation index; T90, proportion of time under 90% oxygen saturation.

Predictive Value of Screening Tools

The predictive performance of the screening tools was presented in Table 3. Compared with participants without insomnia, the STOP-Bang, ESS, and Berlin questionnaires, demonstrated clear trends toward higher sensitivity and lower specificity in screening for all levels of OSA among participants with insomnia. The difference in the predictive value of NoSAS was relatively small between the insomnia and non-insomnia groups.

Among participants with clinical insomnia, the STOP-Bang demonstrated the highest sensitivity [87.4% (95% CI 82.2–91.4) for AHI≥5/h, 93.2% (95% CI 87.6–96.5) for AHI≥15/h, 94.4% (95% CI 86.8–97.9) for AHI≥30/h], while other screening tools had sensitivities less than 80% for all AHI cutoffs. In terms of specificity, the NoSAS was the highest among the four instruments (72.7% (95% CI 62.0–81.4) for AHI≥5/h, 67.7% (95% CI 59.9–74.6) for AHI≥15/h, 62.9% (95% CI 56.2–69.2) for AHI≥30/h). Compared to STOP-Bang, NoSAS brought a net reclassification improvement of 6.8% for AHI≥5/h (P>0.05), 9.7% for AHI≥15/h (P>0.05), and 18.8% for AHI≥30/h (P=0.0017).

The STOP-Bang and NoSAS displayed AUCs higher than 0.7 at all levels of OSA severity (Figure 2). In paired analyses using the DeLong method, the AUCs of STOP-Bang and NoSAS were significantly higher than either the ESS or Berlin questionnaire (p<0.05 in all cases). The comparisons between AUSs of STOP-Bang and NoSAS did not yield significant difference in either AHI≥5/h [0.74 (95% CI, 0.68–0.79) vs 0.73 (95% CI, 0.67–0.80)], AHI≥15/h [0.71 (95% CI, 0.68–0.79) vs 0.72 (95% CI, 0.66–0.77)], or AHI≥30/h [0.73 (95% CI, 0.67–0.79) vs 0.75 (95% CI, 0.70 −0.81)] conditions (P>0.05 in all comparisons).

As the STOP-Bang and NoSAS have an overall advantage among participants with insomnia, the performance of the two tools to predict moderate-severe OSA under each cutoff was determined (Figure 3). The maximum YI for STOP-Bang (0.302, 95% CI 0.190–0.385) was reached at score 4 with sensitivity of 73.7% (95% CI, 65.8–80.5%) and specificity of 56.5% (95% CI, 48.5–64.3%). The maximum YI for NoSAS was reached at score 7 (0.339, 95% CI 0.239–0.416) with sensitivity of 77.2% (95% CI, 69.6–83.7%) and specificity of 56.7% (95% CI, 48.8–64.4%).

Discussion

The current study evaluated on the performance of four commonly used OSA screening tools in the specific population of insomnia. It was noted that the sensitivity and specificity of STOP-Bang, ESS, and Berlin questionnaires had considerable difference between the insomnia and non-insomnia population. Within the insomnia population, STOP-Bang and NoSAS demonstrated an advantage over the ESS or Berlin questionnaire in terms of discriminative power. The conventional cutoff of STOP-Bang score ≥3 provided excellent sensitivity for screening, while analysis of the YI favors a NoSAS cutoff score lower than the originally proposed one.

Table 3 Predictive Value of OSA Screening Tools Among Patients with or Without Insomnia

	Non-insomnia (ISI<15), n=953				Clinical Insomnia (ISI≥15), n=313			
	STOP-Bang (≥3 Points)	**ESS (≥11 Points)**	**Berlin (≥2 Positive Categories)**	**NoSAS (≥8 Points)**	**STOP-Bang (≥3 Points)**	**ESS (≥11 Points)**	**Berlin (≥2 Positive Categories)**	**NoSAS (≥8 Points)**
AHI≥5/h								
Sensitivity (%)	82.2 (79.2–84.9)	44.5 (40.9–48.2)	52.0 (48.3–55.7)	59.4 (55.7–63.0)	87.4 (82.2–91.4)	64.9 (58.2–71.0)	71.3 (64.8–77.0)	56.0 (49.2–62.5)
Specificity (%)	48.3 (41.7–54.9)	63.4 (56.8–69.5)	72.4 (66.1–78.0)	72.4 (66.1–78.0)	34.9 (25.1–46.0)	42.0 (31.8–53.0)	50.0 (39.1–60.9)	72.7 (62.0–81.4)
PPV (%)	83.1 (80.1–85.8)	79.1 (74.7–82.9)	85.4 (81.6–88.5)	87.0 (83.6–89.8)	77.7 (71.9–82.6)	74.1 (67.3–80.0)	78.7 (72.3–84.0)	84.0 (76.9–89.3)
NPV (%)	46.7 (40.3–53.2)	26.9 (23.2–30.8)	32.7 (28.7–37.0)	36.4 (32.1–41.0)	51.7 (38.3–64.9)	31.9 (23.7–41.3)	40.2 (31.0–50.1)	39.3 (31.8–47.2)
AHI≥15/h								
Sensitivity (%)	84.9 (81.4–87.9)	44.6 (40.1–49.1)	54.2 (49.7–58.7)	65.1 (60.7–69.3)	93.2 (87.6–96.5)	67.8 (59.6–75.1)	70.9 (62.8–78.0)	65.1 (56.8–72.6)
Specificity (%)	36.3 (31.9–40.9)	59.5 (54.9–64.0)	62.9 (58.2–67.3)	63.0 (58.4–67.4)	29.8 (23.0–37.6)	41.5 (33.9–49.4)	39.8 (32.2–47.8)	67.7 (59.9–74.6)
PPV (%)	59.2 (55.5–62.8)	54.4 (49.4–59.3)	61.4 (56.7–66.0)	65.7 (61.2–69.8)	55.0 (48.6–61.2)	51.3 (44.1–58.4)	52.0 (44.9–59.0)	64.7 (56.4–72.2)
NPV (%)	68.8 (62.4–74.5)	49.7 (45.5–54.0)	55.8 (51.3–60.1)	62.5 (57.9–66.9)	82.8 (70.1–91.0)	58.6 (49.1–67.6)	59.8 (49.9–69.0)	68.1 (60.3–75.0)
AHI≥30/h								
Sensitivity (%)	86.7 (82.4–90.1)	50.0 (44.4–55.6)	59.0 (53.4–64.3)	70.4 (65.0–75.2)	94.4 (86.8–97.9)	75.3 (64.8–83.5)	77.5 (67.2–85.4)	75.3 (64.8–83.5)
Specificity (%)	31.4 (27.8–35.2)	61.2 (57.3–65.0)	60.6 (56.6–64.4)	58.0 (54.1–61.9)	24.1 (18.7–30.4)	42.0 (35.5–48.7)	39.5 (33.1–46.4)	62.9 (56.2–69.2)
PPV (%)	39.5 (35.9–43.2)	39.9 (35.1–44.9)	43.6 (38.9–48.4)	46.3 (41.9–50.9)	33.5 (27.7–39.7)	34.0 (27.5–41.1)	34.2 (27.8–41.2)	44.7 (36.6–53.0)
NPV (%)	82.1 (76.5–86.6)	70.4 (66.3–74.1)	74.1 (70.0–77.8)	79.2 (75.1–82.7)	91.4 (80.3–96.8)	81.0 (72.5–87.5)	81.3 (72.4–87.9)	86.5 (80.1–91.2)

Abbreviations: ESS, Epworth sleepiness scale; ISI, insomnia severity index; NPV, negative predictive value; OSA, obstructive sleep apnea; PPV, positive predictive value.

https://doi.org/10.2147/NSS.S494804

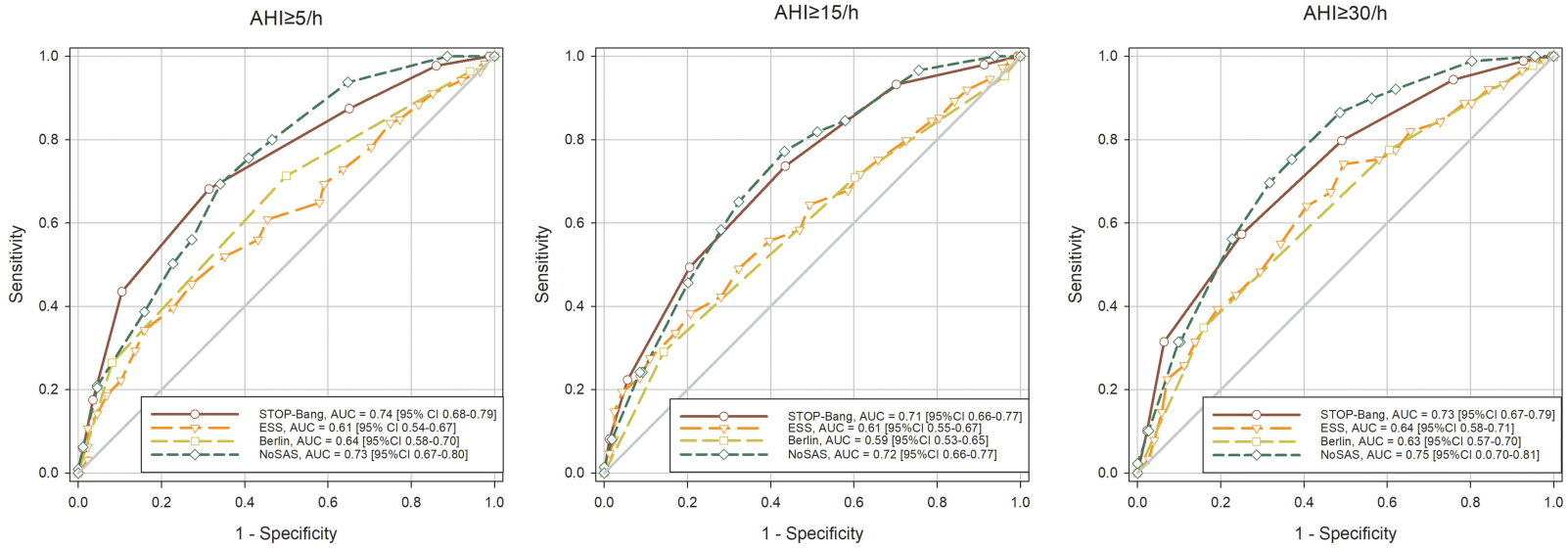

Figure 2 Receiver operating curves of common sleep apnea screening tools in participants with clinical insomnia. Obstructive sleep apnea severity was classified by the AHI. The STOP-Bang questionnaire and NoSAS score had similar AUCs, both performing higher discriminative capacity than the ESS and Berlin questionnaire at all OSA severity.

https://doi.org/10.2147/NSS.S494804

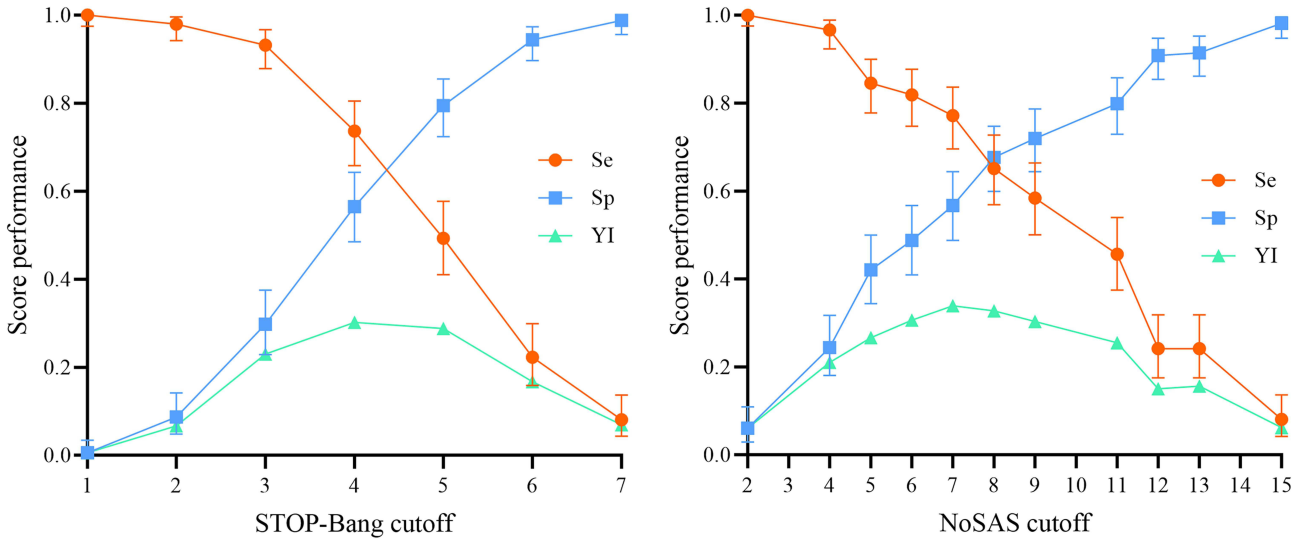

Figure 3 Performance of STOP-Bang and NoSAS to predict moderate-severe obstructive sleep apnea among participants with insomnia at each cutoff. The maximum YI was reached at score 4 for STOP-Bang, and score 7 for NoSAS, respectively. Error bars stand for 95% confidence interval. Data were not shown for NoSAS scores at 1, 14, and 16 as these scores are not reachable by the definition of the tool.
Abbreviations: Se, sensitivity; Sp, specificity; YI, Youden's index.

The clinical application of screening questionnaires focuses on selecting patients who are at high risk of OSA. An essential aspect of an ideal screening tool is the adaptability to both general and specific population.[6] Numerous studies have validated the performance of STOP-Bang, ESS, and Berlin questionnaire in patients referred to sleep clinics for suspected OSA.[24–26] However, when the targeted sample was extended to patients referred to a non-respiratory-based sleep clinic with various sleep disorders, the Berlin questionnaire yielded lower-than-expected performance,[27] suggesting a possibility that other common sleep disorders and/or medications may interfere. The current study demonstrated how the presence of insomnia significantly increases the ESS score and the score of the daytime sleepiness category of Berlin questionnaire. This resulted in a notable change to the performance of ESS and Berlin questionnaire. By contrast, the NoSAS tool, without evaluation of daytime sleepiness, had a similar performance between the insomnia and non-insomnia groups. It should be noted that excessive daytime sleepiness is a highly prevalent condition with various underlying conditions including insufficient sleep, circadian rhythm disorders, central disorders of hypersomnolence, and the use of medications (eg, alcohol, benzodiazepines, opiates).[28] The results from the current study imply that OSA screening tools incorporating measures of daytime function should be used with caution not only in the presence of comorbid insomnia but also in other conditions that increase the likelihood of sleepiness.

In the present study, participants completed all the four screening tools before undergoing PSG, providing a chance to compare the performance of these tools. The results support the use of STOP-Bang and NoSAS, over ESS or Berlin, for superior capability measured by AUCs in the insomnia population. Previous comparative studies conducted in the general population have favored the use of STOP-Bang over Berlin or ESS for higher discriminative capability.[29,30] Likewise, the more recently developed NoSAS, has outperformed Berlin and ESS in a number of reports.[15,20,21] Similar AUCs were observed for STOP-Bang and NoSAS in the current study, which is consistent with two previous sleep clinic-based studies.[15,31] It has also been confirmed that the established correlation between STOP-Bang and NoSAS scores with the severity of OSA[32,33] may extend to individuals with insomnia. Both the STOP-Bang and NoSAS can be easily calculated based on self-reports and simple physical examination, enabling convenient assessment in outpatient settings.

In practice, individuals at risk are selected by screening tools through certain cutoffs. A STOP-Bang score of 3 or more provides high sensitivity with acceptable specificity in most conditions and has been adopted in various populations (eg, general, surgical, sleep clinic, commercial drivers) across different geographic regions.[34,35] However, when the demographic features of a certain population directly influence the STOP-Bang score, as in the case of obesity, a higher cutoff might improve the balance of sensitivity and specificity.[36] In the current study, although STOP-Bang score cutoff

of 4 maximized the YI among the insomnia population, the sensitivity for detecting moderate-severe OSA was below 80%. Considering the application scenario of the questionnaire, we recommend the conventional cutoff score of 3, which retained high sensitivity across the entire spectrum of OSA severity. For NoSAS, the presence of insomnia had little effect on its performance. However, the proposed cutoff of 8 was associated with poor sensitivity and NPV (both below 70%) for AHI≥15/h in the current study, and the cutoff of 7 reached a higher YI with a sensitivity of 77.2%. Notably, the cutoff of 8 provided sensitivity of 79% and NPV of 90% for AHI≥20/h in the original population-based study conducted in Switzerland.[20] The age distribution might be the main factor contributing to this discrepancy, as only 24% of participants in the current study were over 55 years old (the single age cutoff assigned with 4 points), compared to a mean age of 59 years in the derivation study. By contrast, the performance metrics of NoSAS in the current study is more comparable to a population-based cohort study in Singapore with a similar age distribution.[37] It should also be pointed out that Asian OSA patients are reported with lower BMI,[38,39] and whether ethnic-specific BMI cutoff should be adapted has been debated in the field.[40] Moreover, the study setting may have an impact as the cutoff of 7 was proposed by another sleep clinic-based study to achieve optimal performance.[33] In general, the present study opened up discussion on the issue of appropriate cutoff selection for OSA screening tools in the presence of insomnia. External validation of the proposed cutoffs in population samples with ethnic diversity would be highly valuable.

The current study represents an attempt to evaluate the performance of commonly used OSA screening tools among the insomnia population. The findings provide guidance on the selection of instruments, choice of cutoffs, and interpretation of screening results. The study is limited by the selection strategy and single-center design. The inclusion of referred patients increased the pre-test probability of OSA, and the sample consisted mainly of middle-aged overweight males. Validations in population-based samples with a broader age range and a more balanced sex ratio should be sought in the future.

Conclusion

The performance profile of OSA screening tools consisting daytime function measurement may be altered by the presence of insomnia. Among individuals with insomnia who were referred for sleep studies, the STOP-Bang and NoSAS provided overall superior discrimination performance compared with Berlin questionnaire and ESS. A STOP-Bang score ≥3 may sensitively select individuals at risk. The appropriate cutoff score for NoSAS needs further validation.

Acknowledgments

The authors are grateful to all the participants involved in the current study and would like to acknowledge the technical assistance of Ms. Hairong Zhang and Ms. Lijuan Fan.

Funding

The current study was funded by the National High Level Hospital Clinical Research Funding (2022-PUMCH-A-157) and CAMS Innovation Fund for Medical Sciences (CIFMS) (2021-I2M-C&T-B-013).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Gottlieb DJ, Punjabi NM. Diagnosis and management of obstructive sleep apnea: a review. *JAMA*. 2020;323(14):1389–1400. doi:10.1001/jama.2020.3514
2. 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
3. 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
4. Watson NF. Health care savings: the economic value of diagnostic and therapeutic care for obstructive sleep apnea. *J Clin Sleep Med*. 2016;12(8):1075–1077. doi:10.5664/jcsm.6034
5. Drager LF, Santos RB, Pachito D, Albertini CS, Sert Kuniyoshi FH, Eckeli AL. Inequalities in the access to diagnosis and treatment of obstructive sleep apnea in Brazil: a cross-sectional study. *J Clin Sleep Med*. 2024;20(5):735–742. doi:10.5664/jcsm.10976

6. 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(7):1239–1244. doi:10.5664/jcsm.7232
7. Sweetman A, Lack L, McEvoy RD, et al. Bi-directional relationships between co-morbid insomnia and sleep apnea (COMISA). *Sleep Med Rev*. 2021;60:101519. doi:10.1016/j.smrv.2021.101519
8. Lechat B, Loffler KA, Wallace DM, et al. All-cause mortality in people with co-occurring insomnia symptoms and sleep apnea: analysis of the Wisconsin Sleep Cohort. *Nat Sci Sleep*. 2022;14:1817–1828. doi:10.2147/NSS.S379252
9. Guo J, Redline S, Stone KL, Xiao Y. Redefining comorbid insomnia and sleep apnea: the association of sleep breathing impairment and insomnia with incident diabetes. *Ann Am Thorac Soc*. 2023;20(12):1791–1800. doi:10.1513/AnnalsATS.202302-171OC
10. Zhang Y, Ren R, Lei F, et al. Worldwide and regional prevalence rates of co-occurrence of insomnia and insomnia symptoms with obstructive sleep apnea: a systematic review and meta-analysis. *Sleep Med Rev*. 2019;45.
11. 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
12. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. *Ann Intern Med*. 1999;131(7):485–491. doi:10.7326/0003-4819-131-7-199910050-00002
13. 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
14. Hein M, Lanquart J-P, Loas G, Hubain P, Linkowski P. Prevalence and risk factors of excessive daytime sleepiness in insomnia sufferers: a study with 1311 individuals. *J Psychosom Res*. 2017;103:63–69. doi:10.1016/j.jpsychores.2017.10.004
15. Duarte RLM, Magalhães-da-Silveira FJ, Oliveira-E-Sá TS, Rabahi MF, Mello FCQ, Gozal D. 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
16. Connor RJ. Sample size for testing differences in proportions for the paired-sample design. *Biometrics*. 1987;43(1):207–211. doi:10.2307/2531961
17. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med*. 2001;2 (4):297–307. doi:10.1016/S1389-9457(00)00065-4
18. Luo J, Huang R, Zhong X, Xiao Y, Zhou J. Value of STOP-Bang questionnaire in screening patients with obstructive sleep apnea hypopnea syndrome in sleep disordered breathing clinic. *Chin Med J*. 2014;127(10):1843–1848. doi:10.3760/cma.j.issn.0366-6999.20133003
19. Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea. The Epworth Sleepiness Scale. *Chest*. 1993;103(1):30–36. doi:10.1378/chest.103.1.30
20. 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
21. 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
22. Pencina MJ, D'Agostino RB, Vasan RS, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27(2):157–172. doi:10.1002/sim.2929
23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845. doi:10.2307/2531595
24. Kim B, Lee EM, Chung YS, Kim WS, Lee SA. The utility of three screening questionnaires for obstructive sleep apnea in a sleep clinic setting. *Yonsei Med J*. 2015;56(3):684–690. doi:10.3349/ymj.2015.56.3.684
25. 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
26. Pataka A, Daskalopoulou E, Kalamaras G, Fekete Passa K, Argyropoulou P. Evaluation of five different questionnaires for assessing sleep apnea syndrome in a sleep clinic. *Sleep Med*. 2014;15(7):776–781. doi:10.1016/j.sleep.2014.03.012
27. Ahmadi N, Chung SA, Gibbs A, Shapiro CM. The Berlin questionnaire for sleep apnea in a sleep clinic population: relationship to polysomnographic measurement of respiratory disturbance. *Sleep Breath*. 2008;12(1):39–45. doi:10.1007/s11325-007-0125-y
28. Gandhi KD, Mansukhani MP, Silber MH, Kolla BP. Excessive daytime sleepiness: a clinical review. *Mayo Clin Proc*. 2021;96(5):1288–1301. doi:10.1016/j.mayocp.2020.08.033
29. Chiu H-Y, Chen P-Y, Chuang L-P, 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. doi:10.1016/j.smrv.2016.10.004
30. Luo J, Huang R, Zhong X, Xiao Y, Zhou J. STOP-Bang questionnaire is superior to Epworth sleepiness scales, Berlin questionnaire, and STOP questionnaire in screening obstructive sleep apnea hypopnea syndrome patients. *Chin Med J*. 2014;127(17):3065–3070.
31. Veugen CCAFM, Teunissen EM, den Otter LAS, Kos MP, Stokroos RJ, Copper MP. Prediction of obstructive sleep apnea: comparative performance of three screening instruments on the apnea-hypopnea index and the oxygen desaturation index. *Sleep Breath*. 2021;25 (3):1267–1275. doi:10.1007/s11325-020-02219-6
32. Chung F, Subramanyam R, Liao P, Sasaki E, Shapiro C, Sun Y. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth*. 2012;108(5):768–775. doi:10.1093/bja/aes022
33. 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
34. Pivetta B, Chen L, Nagappa M, et al. Use and performance of the STOP-bang questionnaire for obstructive sleep apnea screening across geographic regions: a systematic review and meta-analysis. *JAMA Network Open*. 2021;4(3):e211009. doi:10.1001/jamanetworkopen.2021.1009
35. Nagappa M, Liao P, Wong J, et al. Validation of the STOP-Bang questionnaire as a screening tool for obstructive sleep apnea among different populations: a systematic review and meta-analysis. *PLoS One*. 2015;10(12):e0143697. doi:10.1371/journal.pone.0143697
36. Chung F, Yang Y, Liao P. Predictive performance of the STOP-Bang score for identifying obstructive sleep apnea in obese patients. *Obes Surg*. 2013;23(12):2050–2057. doi:10.1007/s11695-013-1006-z
37. Tan A, Hong Y, Tan LWL, van Dam RM, Cheung YY, Lee C-H. 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
38. Li KK, Kushida C, Powell NB, Riley RW, Guilleminault C. Obstructive sleep apnea syndrome: a comparison between Far-East Asian and white men. *Laryngoscope*. 2000;110(10 Pt 1):1689–1693. doi:10.1097/00005537-200010000-00022

39. Lee RWW, Vasudavan S, Hui DS, et al. Differences in craniofacial structures and obesity in Caucasian and Chinese patients with obstructive sleep apnea. *Sleep*. 2010;33(8):1075–1080. doi:10.1093/sleep/33.8.1075
40. Xia M, Liu S, Ji N, et al. BMI 35 kg/m2 does not fit everyone: a modified STOP-Bang questionnaire for sleep apnea screening in the Chinese population. *Sleep Breath*. 2018;22(4):1075–1082. doi:10.1007/s11325-017-1610-6