



Effects of age and sex on the performance of the NoSAS score as a screening tool for obstructive sleep apnea: a hospital-based retrospective study in China

Zhigang Zhang¹ · Dan Yang¹ · Haiying Wang¹ · Xinmin Liu¹

Received: 23 April 2020 / Revised: 20 October 2020 / Accepted: 13 November 2020 / Published online: 21 November 2020
© Springer Nature Switzerland AG 2020

Abstract

Objectives The NoSAS score has been shown to be a reliable screening tool for obstructive sleep apnea (OSA) in overall populations. This study aimed to explore the effects of age and sex on the predicting performance of this score.

Methods A retrospective analysis was conducted on 1119 subjects aged ≥ 18 years and with a total sleep time of ≥ 4 h during overnight polysomnography. Discrimination was assessed by using areas under receiver operating characteristic curve (AUCs), while predictive parameters were calculated by using contingency tables.

Results Overall, a NoSAS score of 8 points or higher resulted in sensitivity, specificity, and AUC for predicting an apnea–hypopnea index (AHI) of ≥ 20 events/h of 74%, 36%, and 0.63 (in non-elderly 73%, 46%, and 0.65; in elderly 91%, 17%, and 0.59; in men 85%, 18%, and 0.56; in women 52%, 76%, and 0.71, respectively). The AUCs at all AHI cutoffs were significantly lower in men than in women (all with $p < 0.01$), while the AUCs at AHI cutoff of 5, 15, and 30 events/h were significantly lower in elderly than in non-elderly ($p < 0.01$, 0.05, and 0.05, respectively). In non-elderly, a conventional NoSAS with cutoff of 7 or a modified NoSAS with age cutoff of 50 years provided sensitivity and specificity for predicting an AHI of ≥ 20 events/h of 87%, 37% and 80%, 36%, respectively, with comparable AUCs. In women, a conventional NoSAS with cutoff of 6 or a modified NoSAS with neck circumference cutoff of 35 cm provided sensitivity and specificity for predicting an AHI of ≥ 20 events/h of 85%, 39% and 79%, 52%, respectively, with comparable AUCs.

Conclusions NoSAS score has better discrimination but lower sensitivity for predicting OSA in non-elderly and women than in their counterparts. Age- and sex-specific cutoff values reverse this imbalance. Our results underline the preference of age- and sex-specific cutoff values and the need for better age- and sex-specific screening algorithms.

Keywords NoSAS score · Obstructive sleep apnea · Screening · Questionnaire

Introduction

Obstructive sleep apnea (OSA) is the most common type of sleep-disordered breathing (SDB) affecting 9–25% of general adult population [1] and is characterized by recurrent obstruction of the pharyngeal airway during sleep, nocturnal hypoxemia, and excessive daytime sleepiness. Several population-based studies have reported a strong and independent association of OSA with hypertension, stroke, myocardial

infarction, diabetes, long-term cognitive impairment, and increased all-cause mortality [2–5]. Despite the substantial burden of this disease, up to 80% of individuals with moderate-to-severe OSA remain undiagnosed [6].

Although overnight polysomnography (PSG) is the gold standard for diagnosing the presence and severity of OSA, its high expense, relative inaccessibility, and time consumption can delay the diagnosis and treatment of OSA. Therefore, a brief and precise screening tool of identifying patients who are at high risk of OSA and triaging them for prompt diagnosis and treatment is clinically relevant, particularly in areas with limited healthcare resources. A number of screening tests have been developed to identify high-risk patients. However, many of these screening tests are complicated and composed of a couple of subjective items, making them inconvenient to use and may increase variability in clinical practice.

✉ Zhigang Zhang
zhigang700@sina.com

¹ Department of Geriatrics, Peking University First Hospital, Beijing 100034, China

The NoSAS score is a new screening tool based on five items (neck circumference, obesity, snoring, age, and sex), which was developed in a population-based cohort of 2121 subjects in Switzerland (HypnoLaus) and subsequently validated in a Brazilian cohort of 1042 subjects (EPISONO) [7]. Using a definition of SDB as apnea–hypopnea index (AHI) \geq 20 events/h, the authors found that the NoSAS score had high negative predictive values (NPVs) of 90% and 98% in these cohorts, respectively. Moreover, it outperformed the STOP-Bang and Berlin questionnaires, as evidenced by higher area under the curve (AUC) values [7]. In a multiethnic Asian population-based cohort study, Tan et al. [8] found that the NoSAS score performed similarly to the STOP-Bang and Berlin questionnaires. All three scores had high NPVs in ruling out severe SDB. In a hospital-based retrospective study in China, Hong et al. [9] validated that the NoSAS score performed significantly better than the STOP, STOP-Bang, and ESS, and at par with the Berlin questionnaire for SDB screening. The variation in the predictive parameters among different studies may be due to the difference in sample size and discrepancies of demographic and anthropometric parameters as well as associated comorbidities of the recruited subjects. Given the important age- and sex-relevant differences in prevalence, anthropometry, and clinical presentations of OSA, we hypothesize that there is an age- or sex-specific difference in the performance of NoSAS score for predicting OSA with different levels of severity in clinical population. The objective of the present study was to investigate and compare the predictive parameters of NoSAS score among subjects with different age or sex according to the severity of OSA and validate its values as a screening tool in different populations. Additionally, we aimed to determine age- or sex-specific cut-off values to predict OSA in these populations.

Methods

Participants

This was a cross-sectional retrospective study of consecutive subjects suspected of having SDB and referred for overnight PSG at Peking University First Hospital between January 2006 and September 2017. Inclusion criteria were (a) aged \geq 18 years, (b) total sleep time \geq 4 h during PSG, and (c) having complete clinical information associated with NoSAS score. Patients previously diagnosed with OSA, those undergoing noninvasive ventilation pressure titration during PSG, additional diagnoses obtained throughout PSG (like central sleep apnea syndromes and sleep-related hypoventilation disorders), and all cases with technically inadequate PSG were excluded from this study. Our study protocol was approved by the Ethics Committee of Peking University First Hospital and waived the patient consent requirement.

Data extraction

The following data which were collected by a technician prior to PSG was extracted from medical records: neck circumference (NC), body mass index (BMI), snoring, age, and sex. The participants' NoSAS scores were retrospectively calculated from the available data. The NoSAS score allocates 4 points for having a NC $>$ 40 cm, 3 points for having a BMI of 25–30 kg/m², 5 points for having a BMI of 30 kg/m² or more, 2 points for snoring, 4 points for being older than 55 years, and 2 points for being male. Subjects with 8 points or higher are at high risk of OSA [7].

Sleep study

All sleep studies were conducted either on a GraeL diagnostic sleep system (Compumedics, Aus.) or on an EMBLA N7000 digital system (Embla Systems Inc., Broomfield, CO, USA). The following were performed: electroencephalography, electrooculography, electromyography, snore detection, airflow monitoring, respiratory effort assessment, pulse oximetry, electrocardiography, and body position monitoring. PSG records before the year of 2012 were rescored manually by a certified technician and subsequently reviewed by a sleep physician and interpreted in accordance with American Academy of Sleep Medicine criteria [10] as the rest of the records, which define apnea as a \geq 90% reduction in airflow from baseline for \geq 10 s and hypopnea as a \geq 30% reduction in airflow from baseline for \geq 10 s, accompanied by desaturation \geq 3% or an arousal. Both of the technician and the physician were blinded to NoSAS results. The apnea index (AI), hypopnea index (HI), and AHI were calculated by determining the total number of apnea, hypopnea, and apnea plus hypopnea per hour of sleep, respectively. The diagnosis of OSA was based on an AHI \geq 5 events/h with more than 50% of all respiratory events obstructive. The severity of OSA was classified as follows: mild ($5 \leq$ AHI $<$ 15 events/h), moderate ($15 \leq$ AHI $<$ 30 events/h), and severe (AHI \geq 30 events/h). We defined clinically significant OSA as an AHI \geq 20 events/h, according to the NoSAS score study [7]. Indices of nocturnal hypoxemia were the following: mean SpO₂, percentage of recording time with SpO₂ below 90%, minimal SpO₂ value recorded during sleep, and oxygen desaturation index (ODI), which was defined as the number of scored desaturations \geq 3% per hour of sleep.

Statistical analysis

Statistical analysis was performed with Stata, version 14.0 (College Station, TX, USA). Data are presented as mean \pm standard deviation or median with interquartile range (IQR) as appropriate for continuous variables, and as number (percentage) for categorical variables. Groups were compared

with the chi-square test or Fisher exact test (for dichotomous variables), Student’s *t* test (for normally distributed continuous variables), or the Mann–Whitney *U* test (for non-normally distributed continuous variables). Discrimination at six AHI thresholds (5, 10, 15, 20, 25, and 30 events/h) was estimated from the AUC obtained by receiver operator characteristic (ROC) curves and compared between subpopulations (men vs. women, elderly vs. non-elderly), which may range from 0.5 (no discrimination) to 1.0 (perfect discrimination). An AUC > 0.7 was considered clinically significant. Using 2 × 2 contingency tables, the following parameters were calculated: sensitivity, specificity, likelihood ratio positive (LRP), likelihood ratio negative (LRN), accuracy (rate of correct classification), and diagnostic odds ratio (DOR). All statistical tests were two-sided, and values of *p* < 0.05 were considered statistically significant.

Results

Participant characteristics

From the 1627 subjects that underwent overnight PSG, 1119 (68.8%) were included in the study. The exclusions consisted of 103 with a total sleep time < 4 h, 5 with an age younger than 18 years, 186 with a previous diagnosis of OSA and/or

noninvasive ventilation pressure titration during PSG, 166 with incomplete information and/or technical error during data collection, and 48 with additional diagnoses obtained throughout PSG (Fig. 1). The excluded subjects were not different from the study population in age, sex, NC, snoring, and BMI. For the study population, the median age, BMI, NC, and AHI were 58.0 years (IQR 49.0, 71.0), 26.7 kg/m² (IQR 24.2, 29.1), 40.0 cm (IQR 38.0, 43.0), and 35.6 events/h (IQR 21.4, 55.9), respectively. There were 893 (79.8%) men and 226 (20.2%) women. Compared with men, women were older and had lower proportion of snorer; lower median BMI, NC, AHI, and AI; and less severe desaturation, but higher median HI (all with *p* < 0.05). There were 403 (36.0%) elderly participants (aged ≥ 65 years) and 716 (64.0%) non-elderly participants (aged < 65 years). Compared with the non-elderly, the elderly had lower proportion of men and snorer; lower median BMI, NC, and AI; and less severe desaturation, but higher median HI (all with *p* < 0.05) (Table 1).

Prevalence of OSA

Mild OSA was present in 13.8%, 14.2%, 12.9%, 11.8%, and 21.7% of all, non-elderly, elderly, male, and female participants, respectively. Moderate OSA was present in 24.0%, 22.9%, 26.1%, 24.4%, and 22.6% of all, non-elderly, elderly, male, and female participants, respectively. Severe OSA was

Fig. 1 Study protocol and flow diagram. OSA, obstructive sleep apnea; PSG, polysomnography; TST, total sleep time

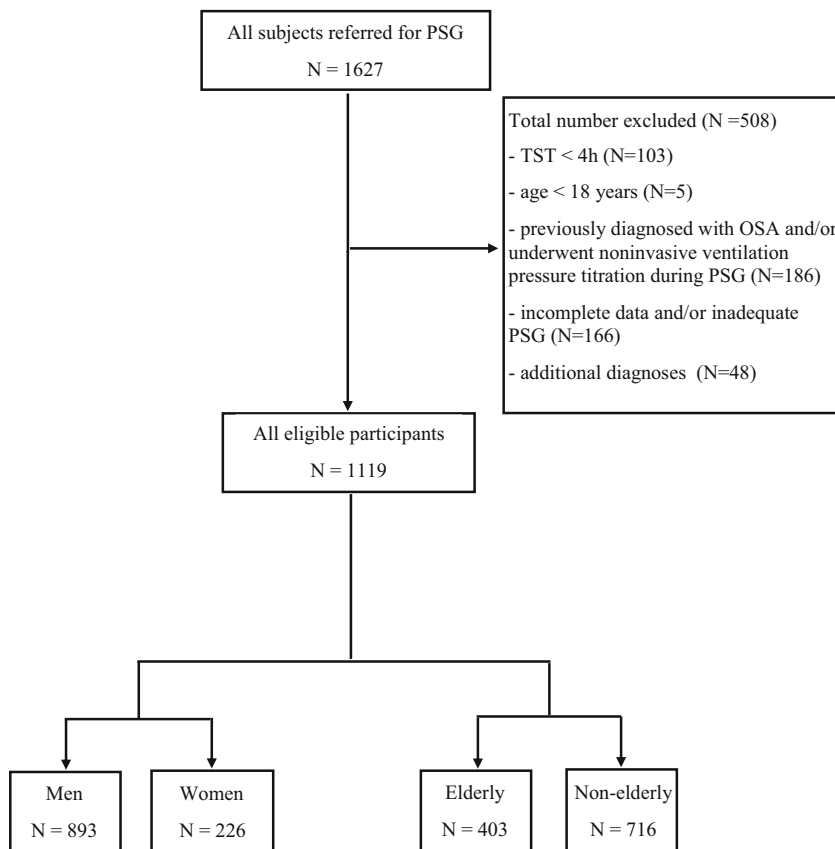


Table 1 Demographic, anthropometric, and sleep data in different populations

	All (<i>n</i> = 1119)	Non-elderly (<i>n</i> = 716)	Elderly (<i>n</i> = 403)	Men (<i>n</i> = 893)	Women (<i>n</i> = 226)
Men, <i>n</i> (%)	893 (79.8)	592 (82.7)	301 (74.7)**	893 (100.0)	0 (0.0)
Age (year) ^Δ	58.0 (49.0, 71.0)	52.0 (44.0, 58.0)	75.0 (70.0, 80.0)**	58.0 (47.0, 71.0)	63.0 (54.0, 72.0) ^{††}
BMI (kg/m ²) ^Δ	26.7 (24.2, 29.1)	27.4 (25.3, 29.4)	25.4 (23.4, 27.7)**	26.9 (24.7, 29.2)	25.8 (23.1, 28.2) ^{††}
NC (cm) ^Δ	40.0 (38.0, 43.0)	41.0 (39.0, 43.0)	40.0 (37.0, 42.0)**	41.0 (39.5, 43.0)	36.3 (35.0, 38.5) ^{††}
Snoring, <i>n</i> (%)	1079 (96.4)	700 (97.8)	379 (94.0)**	875 (98.0)	204 (90.3) ^{††}
AHI (events/h) ^Δ	35.6 (21.4, 55.9)	36.3 (21.4, 58.2)	34.2 (21.5, 51.3)	37.0 (23.0, 56.6)	30.4 (14.6, 50.0) ^{††}
AI (events/h) ^Δ	13.5 (3.7, 32.2)	14.3 (3.9, 36.5)	11.6 (3.5, 28.5)*	15.2 (4.5, 34.8)	7.4 (1.8, 19.0) ^{††}
HI (events/h) ^Δ	16.1 (9.7, 24.1)	15.6 (9.4, 23.4)	17.4 (10.5, 25.5)*	15.8 (9.6, 23.4)	17.4 (10.5, 27.2) [†]
ODI (events/h) ^Δ	25.3 (10.5, 47.3)	26.2 (10.5, 50.7)	23.3 (10.7, 41.3)	26.3 (11.9, 48.2)	19.6 (7.7, 41.8) ^{††}
MSpO ₂ (%)	94.6 ± 2.7	94.6 ± 2.8	94.5 ± 2.4*	94.5 ± 2.7	94.9 ± 2.5 ^{††}
LSpO ₂ (%) ^Δ	83.0 (77.0, 87.0)	82.0 (73.4, 87.0)	84.0 (79.0, 87.0)**	82.0 (76.0, 87.0)	84.0 (78.0, 88.0) [†]
%PO ₂ < 90% ^Δ	1.9 (0.2, 7.5)	2.1 (0.2, 9.6)	1.4 (0.2, 5.3)	2.3 (0.3, 8.6)	0.7 (0.1, 5.8) ^{††}
No OSA, <i>n</i> (%)	27 (2.4%)	22 (3.1%)	5 (1.2%)	18 (2.0%)	9 (4.0%)
Mild OSA, <i>n</i> (%)	154 (13.8%)	102 (14.2%)	52 (12.9%)	105 (11.8%)	49 (21.7%) ^{††}
Moderate OSA, <i>n</i> (%)	269 (24.0%)	164 (22.9%)	105 (26.1%)	218 (24.4%)	51 (22.6%)
Severe OSA, <i>n</i> (%)	669 (59.8%)	428 (59.8%)	241 (59.8%)	552 (61.8%)	117 (51.8%) ^{††}

^Δ Data are presented as median (interquartile range)

*Compared with non-elderly, *p* < 0.05

**Compared with non-elderly, *p* < 0.01

[†] Compared with men, *p* < 0.05

^{††} Compared with men, *p* < 0.01

AHI, apnea–hypopnea index; AI, apnea index; BMI, body mass index; HI, hypopnea index; LSpO₂, lowest oxygen saturation; MSpO₂, mean oxygen saturation; NC, neck circumference; %PO₂ < 90%, percentage of total recording time with SpO₂ < 90%

present in 59.8%, 59.8%, 59.8%, 61.8%, and 51.8% of all, non-elderly, elderly, male, and female participants, respectively (Table 1). There was no statistical difference in prevalence of OSA between non-elderly and elderly. The prevalence of AHI ≥ 15, ≥ 20, and ≥ 30 events/h was significantly higher in men than in women (86.2% vs. 74.3%, 80.2% vs. 65.0%, 61.8% vs. 51.8%, respectively, all with *p* < 0.01).

Diagnostic value

Table 2 shows the demographic data and characteristics of the different populations classified by NoSAS score into high- and low-risk groups. NoSAS score was able to distinguish participants who were men, older, snorer, and with higher BMI and NC which allowed higher AHI, higher ODI, and more severe desaturation to be identified as were seen in the PSG data. But NoSAS score did not show significant differences in the lowest oxygen saturation between high- and low-risk groups in the elderly and men according to our PSG data and did not show significant difference in age between high- and low-risk groups in the elderly.

Overall, a NoSAS score of 8 points or higher resulted in sensitivity, specificity, accuracy, and AUC for the prediction of an AHI of ≥ 20 events/h of 79.4%, 35.9%, 69.4%, and 0.63

(in non-elderly 72.9%, 46.4%, 66.8%, and 0.65; in elderly 90.7%, 16.7%, 74.2%, and 0.59; in men 84.9%, 18.1%, 71.7%, and 0.56; in women 52.4%, 76.0%, 60.6%, and 0.71, respectively) (Table 3, Fig. 2). We also calculated the performance of the NoSAS score using AHI cutoffs of 5, 10, 15, 25, and 30 events/h. The AUCs of the NoSAS score at all AHI cutoff values were significantly lower in men than in women (all with *p* < 0.01), while the AUCs at AHI cutoff values of 5, 15, and 30 events/h were significantly lower in the elderly than in the non-elderly (*p* < 0.01, 0.05, and 0.05, respectively). We found a higher LRP, a higher specificity, but a lower sensitivity with a superior performance in women (Table 3, Fig. 3).

In the non-elderly, a NoSAS cutoff point of 7 provided sensitivity, specificity, and accuracy for the prediction of an AHI of ≥ 20 events/h of 86.7%, 37.4%, and 75.3%, respectively. When using an age cutoff of 50 years in the non-elderly, a NoSAS cutoff point of 8 provided sensitivity, specificity, and accuracy for the prediction of an AHI of ≥ 20 events/h of 79.8%, 35.5%, and 69.6% with a comparable AUC with the conventional age cutoff of 55 years (0.65 (95% CI, 0.60–0.70) vs. 0.65 (95% CI, 0.62–0.69), *p* > 0.05) (Table 4).

In women, a NoSAS cutoff point of 6 provided sensitivity, specificity, and accuracy for the prediction of an

Table 2 Demographic, anthropometric, and sleep data (comparison between low-risk group and high-risk group in different populations)

	All (<i>n</i> = 1119)		Non-elderly (<i>n</i> = 716)		Elderly (<i>n</i> = 403)		Men (<i>n</i> = 893)		Women (<i>n</i> = 226)	
	Low risk (<i>n</i> = 270)	High risk (<i>n</i> = 849)	Low risk (<i>n</i> = 226)	High risk (<i>n</i> = 490)	Low risk (<i>n</i> = 44)	High risk (<i>n</i> = 359)	Low risk (<i>n</i> = 140)	High risk (<i>n</i> = 753)	Low risk (<i>n</i> = 130)	High risk (<i>n</i> = 96)
Men, <i>n</i> (%)	301 (74.7)	592** (82.7)	140 (61.9)	452** (92.2)	0 (0.0)	301** (83.8)	140 (100.0)	753 (100.0)	0 (0.0)	0 (0.0)
Age ^Δ (year)	51.0 (44.0, 56.0)	61.0** (51.0, 73.0)	49.0 (43.0, 53.0)	53.0** (45.0, 59.0)	76.0 (71.5, 79.5)	74.0 (70.0, 80.0)	47.0 (41.0, 52.0)	60.0** (50.0, 73.0)	58.0 (50.0, 72.0)	67.0** (61.0, 72.0)
BMI ^Δ (kg/m ²)	24.5 (22.8, 26.9)	27.2** (25.2, 29.7)	25.2 (23.4, 27.4)	28.1** (26.3, 30.7)	22.3 (21.2, 23.5)	26.0 ** (23.9, 28.0)	26.0 (24.2, 27.7)	27.1** (24.8, 29.4)	23.4 (21.9, 25.0)	28.3** (26.6, 31.1)
NC ^Δ (cm)	38.0 (35.0, 40.0)	42.0** (40.0, 43.0)	38.0 (36, 40)	42.0** (41.0, 44.0)	35.3 (33.0, 37.0)	40.0** (38.0, 42.0)	39.0 (38.0, 40.0)	42.0** (40.0, 44.0)	36.0 (33.0, 37.0)	38.0** (36.0, 41.0)
Snoring, <i>n</i> (%)	244 (90.4)	835** (98.4)	215 (95.1)	485** (99.0)	36 (81.8)	343** (95.5)	130 (92.9)	745** (98.9)	112 (86.2)	92* (95.8)
AHI ^Δ (events/h)	28.7 (14.1, 43.6)	38.5** (23.4, 57.9)	28.7 (13.7, 44.4)	41.2** (23.8, 62.3)	28.9 (15.1, 40.0)	35.8* (22.2, 51.7)	32.1 (21.4, 49.2)	37.7* (23.3, 57.8)	22.6 (11.2, 37.4)	40.5** (24.7, 59.7)
AI ^Δ (events/h)	8.9 (1.7, 20.7)	15.6** (4.5, 35.3)	8.6 (1.3, 13.5)	13.2** (4.1, 29.5)	8.9 (2.2, 22.1)	17.6** (4.9, 42.6)	13.1 (4.4, 27.7)	16.5 (4.5, 36.7)	4.5 (0.8, 13.6)	11.4** (4.1, 28.8)
HI ^Δ (events/h)	15.2 (8.7, 22.5)	16.2 (10.1, 25.1)	15.8 (11.8, 23.2)	17.4 (10.3, 25.7)	15.0 (8.6, 22.2)	15.9 (9.7, 24.1)	15.8 (9.9, 22.3)	15.8 (9.6, 23.5)	14.2 (7.9, 22.8)	22.5** (13.5, 31.5)
ODI ^Δ (events/h)	18.5 (7.1, 34.1)	28.0** (12.4, 50.6)	18.8 (6.5, 35.0)	29.7** (13.6, 55.7)	16.0 (7.9, 28.9)	24.8* (11.0, 42.7)	23.9 (9.3, 39.8)	27.5* (12.3, 50.5)	13.0 (5.2, 27.9)	33.4** (13.1, 51.9)
MSPo ₂ (%)	95.5 ± 2.1	94.3 ± 2.8**	95.5 ± 2.1	94.2 ± 2.9**	95.4 ± 1.8	94.4 ± 2.4**	95.2 ± 2.2	94.4 ± 2.8 **	95.8 ± 2.0	93.7 ± 2.6**
LSpO ₂ ^Δ (%)	84.0 (78.0, 89.0)	82.4** (76.0, 87.0)	83.7 (78.0, 89.0)	81.2** (72.0, 87.0)	86.0 (81.8, 89.0)	84.0 (79.0, 87.0)	81.2 (77.0, 88.0)	83.0 (76.0, 87.0)	86.0 (81.0, 90.0)	82.0** (76.0, 85.8)
%PO ₂ < 90(%) ^Δ	0.7 (0, 3.9)	2.4** (0.4, 10.8)	0.9 (0, 4.3)	2.9** (0.5, 13.3)	0.5 (0.1, 1.4)	1.8** (0.3, 5.6)	1.7 (0.1, 5.6)	2.4* (0.4, 9.8)	0.2 (0, 2.0)	2.6** (0.5, 14.4)

^Δ Data are presented as median (interquartile range)

*Compared with low-risk population, *p* < 0.05

**Compared with low-risk population, *p* < 0.01

AHI, apnea-hypopnea index; AI, apnea index; BMI, body mass index; HI, hypopnea index; LSpO₂, lowest oxygen saturation; MSPo₂, mean oxygen saturation; NC, neck circumference; %PO₂ < 90%, percentage of total recording time with SpO₂ < 90%

Table 3 Performance of NoSAS score of 8 points in different populations at different AHI cutoff values

Populations	AUC (95% CI)	Sensitivity (%)	Specificity (%)	Accuracy (%)	LRP	LRN	DOR
AHI \geq 5 events/h							
All	0.74 (0.72–0.77)	76.7	59.3	76.3	1.88	0.39	4.79
Non-elderly	0.81 (0.78–0.84)	69.7	72.7	69.8	2.56	0.42	6.15
Elderly	0.38 (0.33–0.43)**	88.9	0.0	87.8	0.89	-	-
Men	0.68 (0.64–0.71)	84.8	38.9	83.9	1.39	0.39	3.55
Women	0.92 (0.88–0.95) ^{††}	44.2	100.0	46.5	-	0.56	-
AHI \geq 10 events/h							
All	0.68 (0.65–0.71)	77.9	48.8	75.7	1.52	0.45	3.38
Non-elderly	0.71 (0.68–0.74)	71.1	61.0	70.3	1.82	0.47	3.87
Elderly	0.60 (0.55–0.65)	89.7	20.0	85.4	1.12	0.52	2.15
Men	0.61 (0.58–0.64)	85.1	26.3	81.3	1.15	0.57	2.02
Women	0.83 (0.77–0.87) ^{††}	47.7	96.3	53.5	12.9	0.54	23.87
AHI \geq 15 events/h							
All	0.66 (0.63–0.69)	79.1	40.9	72.9	1.34	0.51	2.62
Non-elderly	0.69 (0.65–0.72)	72.6	51.6	69.0	1.50	0.53	2.83
Elderly	0.59 (0.54–0.64)*	90.2	17.5	79.9	1.09	0.56	1.95
Men	0.60 (0.56–0.63)	85.3	22.0	76.6	1.09	0.67	1.63
Women	0.75 (0.69–0.80) ^{††}	50.6	81.0	58.4	2.67	0.61	4.37
AHI \geq 20 events/h							
All	0.63 (0.60–0.66)	79.4	35.9	69.4	1.24	0.57	2.16
Non-elderly	0.65 (0.62–0.69)	72.9	46.4	66.8	1.36	0.58	2.33
Elderly	0.59 (0.54–0.63)	90.7	16.7	74.2	1.09	0.56	1.96
Men	0.56 (0.53–0.60)	84.9	18.1	71.7	1.04	0.83	1.24
Women	0.71 (0.65–0.77) ^{††}	52.4	76.0	60.6	2.18	0.63	3.47
AHI \geq 25 events/h							
All	0.61 (0.58–0.64)	79.6	32.2	64.6	1.17	0.63	1.86
Non-elderly	0.63 (0.60–0.67)	73.4	42.1	63.4	1.27	0.63	2.02
Elderly	0.58 (0.53–0.63)	90.6	14.3	66.8	1.06	0.66	1.61
Men	0.57 (0.53–0.60)	85.0	17.2	65.2	1.03	0.87	1.18
Women	0.71 (0.65–0.77) ^{††}	54.1	74.2	62.4	2.10	0.62	3.39
AHI \geq 30 events/h							
All	0.61 (0.58–0.64)	80.9	31.6	61.0	1.18	0.61	1.95
Non-elderly	0.64 (0.60–0.67)	75.0	41.3	61.5	1.28	0.61	2.11
Elderly	0.56 (0.51–0.61)*	91.3	14.2	60.3	1.06	0.61	1.73
Men	0.57 (0.54–0.60)	85.9	18.2	60.0	1.05	0.78	1.35
Women	0.71 (0.64–0.77) ^{††}	57.3	73.4	65.0	2.15	0.58	3.70

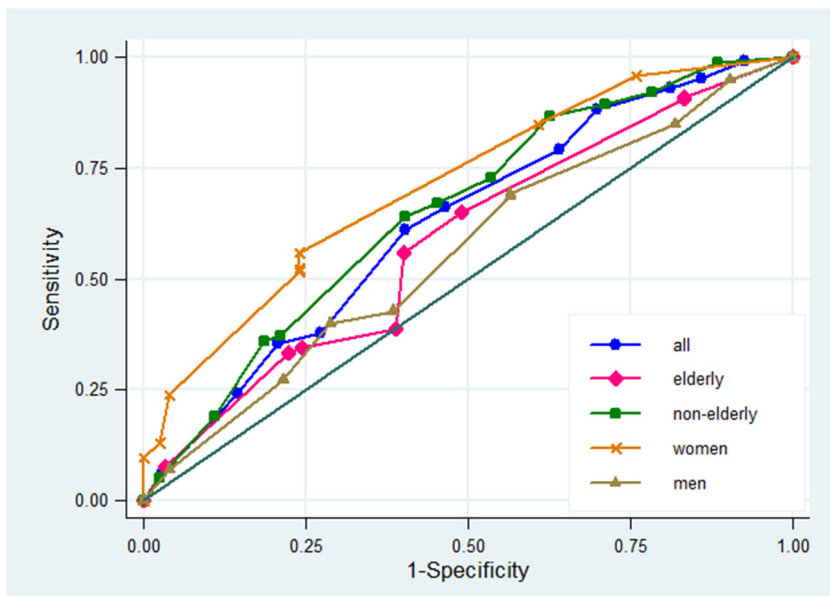
*Compared with non-elderly, $p < 0.05$ **Compared with non-elderly, $p < 0.01$ † Compared with men, $p < 0.05$ †† Compared with men, $p < 0.01$

AHI, apnea–hypopnea index; AUC, area under the curve; CI, confidence interval; DOR, diagnostic odds ratio; LRP, likelihood ratio positive; LRN, likelihood ratio negative

AHI of ≥ 20 events/h of 85.0%, 39.2%, and 69.0%, respectively. When using a NC cutoff of 35 cm in women, a NoSAS cutoff point of 8 provided sensitivity, specificity, and accuracy for the prediction of an AHI of ≥ 20

events/h of 78.9%, 51.9%, and 69.5% with a comparable AUC with the conventional NC cutoff of 40 cm (0.75 (95% CI, 0.69–0.81) vs. 0.71 (95% CI, 0.65–0.77), $p > 0.05$) (Table 5).

Fig. 2 Receiver operating characteristic curves of the NoSAS score at AHI cutoff of ≥ 20 events/h in different populations. At an AHI cutoff of ≥ 20 events/h, the diagnostic performance of the NoSAS score was better in women than in men. AHI, apnea–hypopnea index



Discussion

The NoSAS score is a simple, efficient, and easy-to-implement score that limits the number of subjective variables to only snoring. In this study, we retrospectively analyzed the value of the NoSAS score in subjects with suspected OSA and conducted PSG monitoring in our sleep center. We observed clear age- and sex-specific differences in the performances of NoSAS with significantly better discrimination but lower sensitivity for predicting OSA in the non-elderly and women. Age- and sex-specific cutoff values reverse this imbalance.

Analysis of a questionnaire’s performance in specific populations can provide clinicians with a set of predictive parameters for different levels of OSA severity, which can be used as

a crucial guide for diagnostic and therapeutic decisions. Using a similar AHI cutoff of ≥ 20 events/h, we found that the sensitivity of the NoSAS score (79.4%) was similar to those reported in the HypnoLaus (79.0%) [7] and EPISONO (85.0%) [7] cohorts and another hospital-based cohort in China (74.8%) [9], and higher than that reported in a multiethnic Asian population-based cohort (69.4%) [8]. However, the specificity (35.9%) was lower than the other four cohorts (69.0%, 77.0%, 54.1%, and 78.2%, respectively). The AUC of the NoSAS score (0.63) in our cohort was also slightly lower than the other four cohorts (0.74, 0.81, 0.71, and 0.74, respectively). Apart from differences in sampling methodology, type of sleep studies employed, and scoring criteria used, we suppose that the reduced AUC and specificity of the

Fig. 3 Performance of the NoSAS score in different populations. The AUC of the NoSAS score was lower in the elderly and men than in their counterparts. *Compared with non-elderly, $p < 0.05$. **Compared with non-elderly, $p < 0.01$. ^^Compared with men, $p < 0.01$. AHI, apnea–hypopnea index; AUC, area under the curve

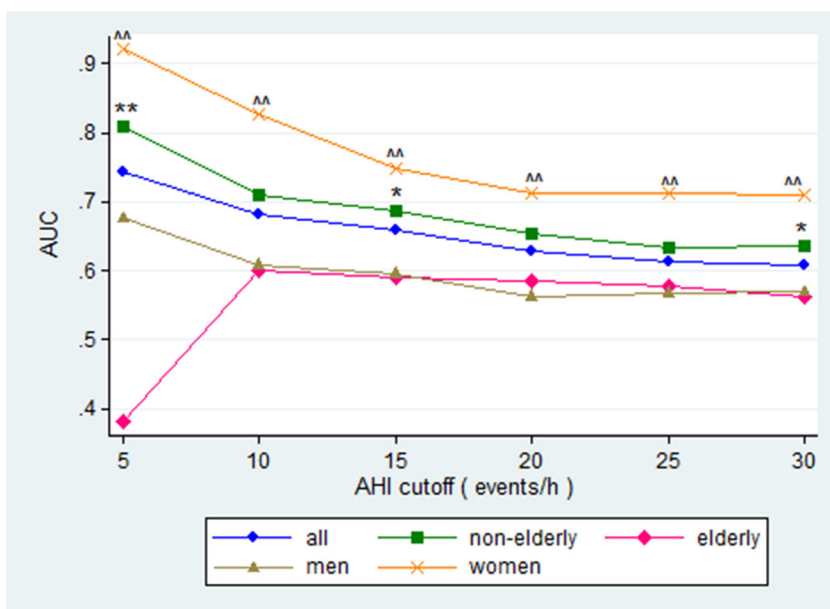


Table 4 Performance of modified NoSAS score of 8 points with age cutoff of 50 years in non-elderly

Cutoff (events/h)	AUC (95% CI)	Sensitivity (%)	Specificity (%)	Accuracy (%)	LRP	LRN	DOR
AHI \geq 5	0.75 (0.66–0.85)	77.1	50.0	76.3	1.54	0.46	3.36
AHI \geq 10	0.69 (0.61–0.76)	77.8	40.7	74.7	1.31	0.55	2.40
AHI \geq 15	0.68 (0.63–0.73)	79.1	37.1	71.8	1.26	0.56	2.23
AHI \geq 20	0.65 (0.60–0.70)	79.8	35.5	69.6	1.24	0.57	2.18
AHI \geq 25	0.63 (0.59–0.68)	80.3	32.5	65.1	1.19	0.61	1.96
AHI \geq 30	0.63 (0.59–0.68)	81.5	31.6	61.5	1.19	0.58	2.04

AHI, apnea–hypopnea index; AUC, area under the curve; CI, confidence interval; DOR, diagnostic odds ratio; LRP, likelihood ratio positive; LRN, likelihood ratio negative

NoSAS score in our cohort could be partly accounted by the differences in the constitution of race, age, and sex. Compared with Caucasian and Latino, the pathogenesis of OSA in Chinese could be related to other factors such as craniofacial restriction, neuromuscular control of upper airway, or arousal threshold besides obesity [11, 12]. In our cohort, the proportion of men (79.8%) is much higher than most of the other four cohorts (48.0%, 45.0%, 77.9%, and 50.4% respectively). And the participants (59 ± 15 years) are also much older than most of the other four cohorts (59 ± 11 , 42 ± 14 , 48 ± 14 , and 48 ± 14 years, respectively), which means that there is a much higher proportion of elderly. As our results showed, AUC and specificity of the NoSAS score in the elderly and men were not as good as those in the non-elderly and women. With higher proportions of elderly and men in our cohort, the overall performance of the NoSAS score in screening OSA was barely satisfactory. This hypothesis is supported by another large hospital-based cohort study consisting mainly of men (69.2%) with relatively older participants (mean age of 52.2 years) as well as higher prevalence of overall and severe OSA (97.1% and 64.4%, respectively) [13]. In this study, a NoSAS score of 8 points or higher provided sensitivity and specificity for predicting severe OSA of 90.1% and 29.4% with an AUC of 0.66, which were very close to ours and also lower than the other four cohorts mentioned above.

In this study, we found a higher AUC, a higher specificity, but a lower sensitivity for NoSAS score in women than in men. The sex-specific difference in the performance of the NoSAS score was also found by other researchers in bariatric, morbidly obese, and general populations [14–16]. A recently published population-based cohort study [16] by Bauters et al. using polygraphic data showed that diagnostic performance indices of NoSAS calculated on the overall group (men + women) overestimated the performance in both sexes separately. The sensitivity of NoSAS for an AHI > 15 events/h was acceptable in men (87.1%), but low in women (55.3%). The reverse was true for the specificity (39.9% in men, 87.4% in women) [16]. Our study, similar to previous studies [16–18], demonstrated that OSA was much more prevalent and severe in men than in women, and men have larger NC and BMI than women, whereas women with OSA are generally older than their male counterparts [17–19]. All these differences may cause the disparity in the sensitivity for NoSAS score between men and women. In addition, men generally present with typical symptoms, such as snoring and observed apnea, whereas women are more likely to present with atypical symptoms, such as depression, fatigue, and insomnia [20, 21]. Although NoSAS score limits the number of subjective variables to only snoring compared with other questionnaires, such as STOP-Bang questionnaire, the influence of sex on symptoms associated with OSA like snoring as showed in our study and

Table 5 Performance of modified NoSAS score of 8 points with neck circumference cutoff of 35 cm in women

Cutoff (events/h)	AUC (95% CI)	Sensitivity (%)	Specificity (%)	Accuracy (%)	LRP	LRN	DOR
AHI \geq 5	0.85 (0.75–0.95)	70.1	77.8	70.4	3.15	0.39	8.19
AHI \geq 10	0.79 (0.71–0.87)	72.9	66.7	72.1	2.19	0.41	5.37
AHI \geq 15	0.75 (0.69–0.82)	75.6	53.5	69.9	1.62	0.46	3.56
AHI \geq 20	0.75 (0.69–0.81)	78.9	51.9	69.5	1.64	0.41	4.04
AHI \geq 25	0.75 (0.69–0.81)	80.5	49.5	67.7	1.59	0.40	4.03
AHI \geq 30	0.75 (0.69–0.81)	82.1	46.8	65.0	1.54	0.38	4.02

AHI, apnea–hypopnea index; AUC, area under the curve; CI, confidence interval; DOR, diagnostic odds ratio; LRP, likelihood ratio positive; LRN, likelihood ratio negative

others can still result in the difference in predictive performance between men and women. Thus, a better sex-specific screening algorithm or a sex-specific adjustment of the cutoff values of the scores could help to improve the predictive power in patients with different sex. An ideal OSA screening score should have a high sensitivity to avoid false-negative results, but also be specific enough to avoid referral of low-risk patients for costly and time-consuming sleep recordings. NoSAS score had higher diagnostic accuracy in women; however, its sensitivity was quite low, yielding the highest proportion of missed diagnoses, which was unfortunate for a screening score. With a decreased cutoff value of 6 points, sensitivity was significantly improved, reaching 85.0% for the prediction of an AHI of ≥ 20 events/h with an acceptable specificity of 39.2%. Given the disparity of NC between sexes, sex-specific adjustment of the cutoff values of NC of the score is also a reasonable and practical strategy because the different cutoff values for women and men can be attributed to the differences in the fat distribution and anatomical features between the sexes [17]. The median NC in women of our cohort was 36.3 cm. As our data showed, using a NC cutoff of 35 cm in women could also help to improve the sensitivity (78.9%) for the prediction of an AHI of ≥ 20 events/h with a comparable AUC with the conventional score. These optimal sex-specific NoSAS score and NC cutoff for OSA screening in women have also been validated in the aforementioned population-based study [16]. For men, although the sensitivity of NoSAS score was good enough, there is still much space to improve the diagnostic accuracy.

Studies show that the prevalence of sleep apnea tends to increase with age to approximately 65 and 55 years for women and men respectively, after which it appears to decline [22, 23]. We think that this is the reason why in our study NoSAS score did not show significant difference in age between high- and low-risk groups in the elderly. There were higher sensitivity and lower AUC for NoSAS score in the elderly than in the non-elderly. Meta-regression analysis findings reveal that age and BMI were positively correlated with the pooled sensitivity [24]. Although the elderly had lower BMI and NC, the differences were only mild and the vast majority of the elderly and non-elderly were still lying in the same category according to the associated cutoff values. In addition, with a median age of 52 years, only minority of the non-elderly were older than 55 years and allocated 4 points, meanwhile 4 points were directly assigned to all the elderly. The difference in age may far outweigh the differences in BMI and NC and underlie the observed disparities in sensitivity for the NoSAS score. With a decreased cutoff value of 7 points, sensitivity was significantly improved, reaching 86.7% for the prediction of an AHI of ≥ 20 events/h in the non-elderly with an acceptable specificity of 37.4%. Using a decreased age cutoff value of 50 years as other screening tools, such as STOP-Bang, also improved the sensitivity (79.8%) for the prediction of an AHI of ≥ 20 events/h

in the non-elderly with a comparable AUC with the conventional score. As the prevalence of obesity based on BMI progressively increases until 60 years of age and decreases thereafter [25], the high prevalence of OSA in the elderly could be explained by other factors besides obesity: decreases in ventilator control and muscular endurance, increased upper-airway collapsibility, and sleep fragmentation potentially contribute to the development of OSA in the elderly [26]. A screening algorithm incorporating variables reflecting these possible mechanisms could help to improve the predictive power in the elderly.

One of the strengths of the analysis is that it relies on a large hospital-based cohort consisting of high proportions of the elderly and women, with PSG recordings using up-to-date techniques and manually analyzed according to the current guideline, which probably increases its reliability and relevance for clinical real-world settings. Although our study was a retrospective one, all information was collected before PSG monitoring in order to avoid bias. Furthermore, this is the first study that was effectively designed to investigate age- and sex-specific differences in the performances of NoSAS in Chinese clinical population. There are also limitations that need to be taken into account. First, it involved patients referred to sleep laboratory, which could represent a selection bias and limit its external validity. The prevalence of OSA was 97.6%, and the proportion of severe OSA patients was about 60.4%. So the sensitivity of NoSAS score in our study may be overestimated. However, the comparative results were still reliable and clinically relevant. Second, some subjects were excluded from this study for incomplete data and/or technically inadequate PSG. But the excluded subjects were not different from study population in age, sex, NC, snoring, and BMI.

Conclusions

In subjects referred to sleep laboratory, age- and sex-specific differences in the performance of NoSAS score are present, and their ability to replace diagnostic sleep studies currently appears to be limited, particularly in the elderly and men. If applied, age- or sex-specific cutoff values should be preferred, but there is a need for better age- or sex-specific screening algorithms.

Authors' contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Zhigang Zhang, Dan Yang, and Haiying Wang. The first draft of the manuscript was written by Zhigang Zhang. The draft was reviewed and edited by Zhigang Zhang and Xinmin Liu. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability Not applicable.

Compliance with ethical standards

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

Ethical approval This study protocol was approved by the Ethics Committee of Peking University First Hospital and waived the patient consent requirement.

Consent to participate Not applicable.

Consent for publication Not applicable.

Code availability Not applicable.

References

- Peppard PE, Young T, Barnett JH, Palta M, Hagen EW, Hla KM (2013) Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol* 177:1006–1014. <https://doi.org/10.1093/aje/kws342>
- McNicholas WT, Bonsignore MR (2007) Sleep apnoea as an independent risk factor for cardiovascular disease: current evidence, basic mechanisms and research priorities. *Eur Respir J* 29:156–178. <https://doi.org/10.1183/09031936.00027406>
- Young T, Finn L, Peppard PE, Szklo-Coxe M, Austin D, Nieto FJ, Stubbs R, Hla KM (2008) Sleep disordered breathing and mortality: eighteen-year follow-up of the Wisconsin sleep cohort. *Sleep* 31:1071–1078
- Drager LF, Togeiro SM, Polotsky VY, Lorenzi-Filho G (2013) Obstructive sleep apnea: a cardiometabolic risk in obesity and the metabolic syndrome. *J Am Coll Cardiol* 62:569–576. <https://doi.org/10.1016/j.jacc.2013.05.045>
- Marshall NS, Wong KK, Liu PY, Cullen SR, Knuiman MW, Grunstein RR (2008) Sleep apnea as an independent risk factor for all-cause mortality: the Busselton Health Study. *Sleep* 31:1079–1085
- Young T, Evans L, Finn L, Palta M (1997) Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep* 20:705–706. <https://doi.org/10.1093/sleep/20.9.705>
- 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. [https://doi.org/10.1016/S2213-2600\(16\)30075-3](https://doi.org/10.1016/S2213-2600(16)30075-3)
- 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. <https://doi.org/10.1007/s11325-016-1455-4>
- 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. <https://doi.org/10.5664/jcsm.6930>
- 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 (2012) Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 8:597–619. <https://doi.org/10.5664/jcsm.2172>
- Lee RW, Vasudavan S, Hui DS, Prvan T, Petocz P, Darendeliler MA, Cistulli PA (2010) Differences in craniofacial structures and obesity in Caucasian and Chinese patients with obstructive sleep apnea. *Sleep* 33:1075–1080. <https://doi.org/10.1093/sleep/33.8.1075>
- Li KK, Kushida C, Powell NB, Riley RW, Guilleminault C (2000) Obstructive sleep apnea syndrome: a comparison between Far-East Asian and white men. *Laryngoscope* 110:1689–1693. <https://doi.org/10.1097/00005537-200010000-00022>
- Topirceanu A, Udrescu M, Udrescu L, Ardelean C, Dan R, Reisz D, Mihaicuta S (2018) SAS score: targeting high-specificity for efficient population-wide monitoring of obstructive sleep apnea. *PLoS One* 13:e0202042. <https://doi.org/10.1371/journal.pone.0202042>
- 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. <https://doi.org/10.1007/s11695-018-3222-z>
- Duarte RLM, Mello FC, Magalhaes-da-Silveira FJ, Oliveira-E-Sa TS, Rabahi MF, Gozal D (2019) Comparative performance of screening instruments for obstructive sleep apnea in morbidly obese patients referred to a sleep laboratory: a prospective cross-sectional study. *Sleep Breath* 23:1123–1132. <https://doi.org/10.1007/s11325-019-01791-w>
- Bauters FA, Loof S, Hertegonne KB, Chirinos HA, De Buyzere ML, Rietzschel ER (2020) Sex-specific sleep apnea screening questionnaires: closing the performance gap in women. *Sleep Med* 67:91–98. <https://doi.org/10.1016/j.sleep.2019.10.023>
- Mohsenin V (2003) Effects of gender on upper airway collapsibility and severity of obstructive sleep apnea. *Sleep Med* 4:523–529. [https://doi.org/10.1016/s1389-9457\(03\)00168-0](https://doi.org/10.1016/s1389-9457(03)00168-0)
- Subramanian S, Jayaraman G, Majid H, Aguilar R, Surani S (2012) Influence of gender and anthropometric measures on severity of obstructive sleep apnea. *Sleep Breath* 16:1091–1095. <https://doi.org/10.1007/s11325-011-0607-9>
- Gabbay IE, Lavie P (2012) Age- and gender-related characteristics of obstructive sleep apnea. *Sleep Breath* 16:453–460. <https://doi.org/10.1007/s11325-011-0523-z>
- Basoglu OK, Tasbakan MS (2018) Gender differences in clinical and polysomnographic features of obstructive sleep apnea: a clinical study of 2827 patients. *Sleep Breath* 22:241–249. <https://doi.org/10.1007/s11325-017-1482-9>
- Shah N, Hanna DB, Teng Y, Sotres-Alvarez D, Hall M, Loreda JS, Zee P, Kim M, Yaggi HK, Redline S, Kaplan RC (2016) Sex-specific prediction models for sleep apnea from the Hispanic Community Health Study/Study of Latinos. *Chest* 149:1409–1418. <https://doi.org/10.1016/j.chest.2016.01.013>
- Bixler EO, Vgontzas AN, Lin HM, Ten Have T, Rein J, Vela-Bueno A, Kales A (2001) Prevalence of sleep-disordered breathing in women: effects of gender. *Am J Respir Crit Care Med* 163:608–613. <https://doi.org/10.1164/ajrcm.163.3.9911064>
- Bixler EO, Vgontzas AN, Ten Have T, Tyson K, Kales A (1998) Effects of age on sleep apnea in men: I. Prevalence and severity. *Am J Respir Crit Care Med* 157:144–148. <https://doi.org/10.1164/ajrcm.157.1.9706079>
- Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM (2005) Case-control and two-gate designs in diagnostic accuracy studies. *Clin Chem* 51:1335–1341. <https://doi.org/10.1373/clinchem.2005.048595>
- Mokdad AH, Bowman BA, Ford ES, Vinicor F, Marks JS, Koplan JP (2001) The continuing epidemics of obesity and diabetes in the United States. *JAMA* 286:1195–1200. <https://doi.org/10.1001/jama.286.10.1195>

26. Malhotra A, Huang Y, Fogel R, Lazic S, Pillar G, Jakab M, Kikinis R, White DP (2006) Aging influences on pharyngeal anatomy and physiology: the predisposition to pharyngeal collapse. *Am J Med* 119:72e9–7e14. <https://doi.org/10.1016/j.amjmed.2005.01.077>

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