

Sequential screenings improve prediction in obstructive sleep apnoea severity

^{1,3}Khairul Izzat Bin Hasan *BSc*, ^{1,3}Kheng-Seang Lim *FRCP*, ²Liang-Chye Goh *MS (ORL-HNS)*,
¹Si-Lei Fong *FRCP*, ³Letchumy Praba Ramanaidu *BSc*, ¹Zheng Yang Lee *MBBS*,
¹Chong-Tin Tan *MD*

¹Division of Neurology, Department of Medicine, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia; ²Department of Otorhinolaryngology, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia; ³Neurology Laboratory, Neurology Unit, University Malaya Medical Centre, Kuala Lumpur, Malaysia

Abstract

Objectives: Sleep apnoea syndrome (SAS) poses serious health risks and requires Polysomnography (PSG) for diagnosis. Due to PSG's labour-intensive process and long waiting times, there's a need for predictive models to prioritize severe SAS cases for early PSG. **Methods:** We retrospectively reviewed PSG cases from the University Malaya Medical Centre's ENT clinic (January-December 2023). Data included demographics, anthropometrics, sleep patterns, STOP-BANG questionnaire, Epworth Sleepiness Scale (ESS), and upper airway assessments, analysed for their relationship with AHI. All variables that exhibited statistical significance were categorised into two sets of combined variables: Clinical Examination Score (CES) and ENT Examination Score (EES). **Results:** We studied 201 cases, with an average age of 47.8±15.4 years (range: 16–89). Of these, 125 (62.2%) were male, with a mean AHI of 49.0±33.5. Severe sleep apnoea (AHI≥30) was diagnosed in 127 (63.2%) cases. Significant differences in AHI ($p<0.05$) were found based on gender, snoring, apnoea, nocturia, drooling, BMI>35kg/m², neck circumference > 40cm, Q1 of ESS ($r=0.17$), Q3 of ESS ($r=0.23$), Q6 of ESS ($r=0.17$), total ESS score ($r=0.20$), Modified Mallampati ($r=0.26$), Palatine Tonsil grade ($r=0.27$), Retropalatal grade ($r=0.29$), and Retrolingual grade ($r=0.29$). Hierarchical Multiple Regression analysis explained 33.0% of the variance in AHI, $F(12,188)=7.2$, $p<0.01$. CES had the highest correlation with AHI ($R=0.475$) and AUC (0.735) in ROC analysis, showing high sensitivity (82.7%) and moderate specificity (50.0%); combined with EES, specificity improved to 72.5%.

Conclusion: The CES and EES help prioritize moderate and severe OSA patients for early diagnosis and treatment. This reduces complications, eases healthcare workload, and shortens diagnostic wait times.

Keywords: Sleep apnoea syndrome, obstructive sleep apnoea, Epworth Sleepiness Scale, STOPBANG questionnaire, polysomnography

INTRODUCTION

Sleep apnoea syndrome (SAS) is a prevalent sleep disorder that poses significant health risks and affects a substantial portion of the global population. The patients with SAS have recurring episodes of breathing interruptions, which are called apnoea (total stop breathing) or hypopnea (inadequate volume of breathing).¹ There are two main types of sleep apnoea syndrome: Obstructive sleep apnoea (OSA) and Central sleep apnoea (CSA). The most common type is OSA, when throat muscles relax and block

airflow into the lungs. CSA is a condition in which the brain fails to communicate with the breathing muscles properly. The effects of SAS on people are not only limited to sleep disorders; they also include a range of negative effects, such as excessive daytime sleepiness(EDS), relatively low quality of life, reduced learning skills, and neurocognitive impairment such as decreases in alertness, memory loss, concentration difficulties, and learning abilities.² Long-term untreated sleep apnoea can increase cardiovascular complications, diabetes, and heart failure.³

Address correspondence to: Liang Chye Goh, Department: Department of Otorhinolaryngology, Department of Medicine, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia. Email: gohlc@ummc.edu.my; Kheng Seang Lim, Division of Neurology, Department of Medicine, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia. Email: kslimum@gmail.com

Date of Submission: 30 December 2024; Date of Acceptance: 3 January 2025

<https://doi.org/10.54029/2025vjik>

Prioritizing treatment for severe obstructive sleep apnea (OSA) over mild OSA is essential due to the significant health risks associated with severe cases. Severe apnea is characterized by more frequent and prolonged interruptions in breathing during sleep, leading to increased cardiovascular stress, a higher likelihood of hypertension, stroke, and heart disease.⁴ In contrast, mild apnea may not present immediate life-threatening issues and can often be managed with lifestyle changes such as weight loss, physical exercise, positional therapy, and the use of oral appliances.⁵ As recommended by Yeghiazarians, *et al.*⁶, severe apnea typically requires more urgent treatment options such as continuous positive airway pressure (CPAP) therapy or surgical interventions, while oral appliances may be an option for individuals with mild to moderate OSA or for those who cannot tolerate CPAP. The urgency of addressing severe OSA is crucial, as delaying treatment can lead to exacerbated symptoms, chronic health conditions, and greater healthcare costs in the long run. Thus, timely intervention for severe cases is vital to mitigate risks and enhance overall patient outcomes.

Predicting the severity of sleep apnoea syndrome involves a comprehensive approach, beginning with assessing various risk factors. Demographic information, including age, gender, and family history, provides crucial insights into potential patterns. Elevated Body Mass Index (BMI) serves as a common risk indicator, emphasising the significance of weight in relation to height. Screening questionnaires like the Epworth Sleepiness Scale (ESS) and the STOP-BANG Questionnaire aid in gauging daytime sleepiness and identifying elevated risk of sleep apnoea syndrome.⁷ Observing and gathering self-reported symptoms, such as loud snoring and breathing interruptions, contribute to a diagnosis. Monitoring co-morbidities and conducting clinical assessments of the airway, including the nasal and oral cavity, contribute to a comprehensive predictive model. However, despite the advancements in predictive models, no one accurately predicted the severity of SAS. The STOPBANG questionnaire demonstrates very high sensitivity but low specificity, resulting in a high rate of false positives.⁸ Similarly, the ESS has moderate sensitivity and specificity, which may lead to missed diagnoses of SAS patients.⁷ Due to this, the gold standard for confirmation for SAS remains Polysomnography (PSG), a comprehensive sleep test conducted in a controlled environment that requires at least 8 hours of

test recording and days to analyse. Because of the labour intensity, the waiting time for a PSG in our centre is very long, up to one year. This highlights the challenge of diagnosis using PSG in confirming the presence of the disorder and, subsequently, delay in the treatment.

Sleep apnoea syndrome (SAS) presents a complex and multifaceted challenge to both clinicians and individuals affected. The spectrum of SAS severity varies widely, necessitating understanding the factors contributing to its manifestation. Despite the recognition of sleep apnoea as a public health concern, a comprehensive understanding of the specific factors or combined factors contributing to its severity level remains elusive. This study recognises the limitations of existing prediction tools, as they primarily focus on the presence of SAS, with Polysomnography (PSG) maintaining its gold standard status for grading SAS severity. The reliance on PSG, limited testing centres, and increasing demand create burdensome waiting lists. This situation emphasises the critical need to explore and identify predictors for SAS severity, enabling the development of personalised and efficient management strategies. By focusing on self-reported data, Epworth Sleepiness Scale (ESS), anthropometric measurements, and detailed anatomical assessments of the nasal cavity, oral cavity, and oropharynx, this research aims to untangle the complex interrelationships of factors that will contribute to the specificity and sensitivity in indicating the severity of sleep apnoea syndrome.

METHODS

Study population

We conducted a retrospective cross-sectional study involving a cohort of 201 individuals assessed between January and December 2023, using data from electronic medical records (EMR) of UMMC's database for patients referred to the UMMC sleep laboratory. Participants had initially attended the ENT clinics at UMMC for self-reported data, ESS questionnaires, anthropometric measurements, and upper airway assessments before undergoing Level 1 PSG at the UMMC Sleep Laboratory. We excluded all cases with incomplete data to ensure the accuracy and reliability of the findings. This study was approved by the Medical Ethics Committee of the University Malaya Medical Centre (MREC Number: 2024515-13733).

Data collection

In our study, we retrieved the data of 201 patients: (1) Demographic data (e.g., age, gender, ethnicity, partner); (2) Anthropometric parameters (height, weight, neck circumference, waist circumference, etc.); (3) Previous clinical history (history of hypertension, diabetes, cardiovascular and cerebrovascular disease, neurological disease and other related diseases); (4) Sleep-related breathing and other sleep-related events (e.g., snoring, apnoea, choking, nocturia); (5) Epworth-sleepiness and STOPBANG scales; (6) ENT's upper airway assessment; (7) Apnoea-hypopnea index (AHI).⁹

Apnoea-Hypopnea Index (AHI) value

An Apnoea + Hypopnea index (AHI) is used to quantify and assess the frequency of obstructive events during sleeping and identify the severity of OSA. It is defined as the average number of apnoea and hypopnea per hour. Apnoea was defined by the American Academy of Sleep Medicine (AASM) as a complete cessation of breathing with more than 90% decrease in airflow for a minimum of 10 seconds, while hypopnea is defined as partial cessation of breathing where airflow decreases at least 30% and must be accompanied by oxygen saturation decrement at $\geq 3\%$ or associated with arousal.¹⁰ Based on the AHI value, OSA will be classified as follows: AHI < 5 per hour is normal, ≥ 5 but < 15 per hour is mild, ≥ 15 but < 30 is moderate, and severe if ≥ 30 .¹¹ In this study, level 1 PSG with a minimum of 6 hours of recording were used on every patient and underwent analysis to determine the sleep stages and the respiratory event for AHI. Sleep stages and respiratory event scoring were determined manually by a qualified neurophysiology technologist based on the AASM version 2.4 (2017) manual guidelines.

Epworth Sleepiness Scale (ESS)

The Epworth Sleepiness Scale (ESS) comprises eight questions to assess subjective sleepiness across various everyday scenarios. Each question prompts respondents to rate the likelihood of dozing off on a scale of 0 to 3 points, where 0 indicates "would never doze" and 3 indicates a "high chance of dozing".⁹ Therefore, the total ESS score ranges from 0 to 24, reflecting the cumulative likelihood of dozing across these scenarios. Individuals with an ESS score exceeding 10 were categorised as experiencing

significant daytime sleepiness based on the scale's established criteria.¹²

STOPBANG

The STOP-Bang questionnaire comprises eight questions related to snoring, tiredness, observed apnoeas, hypertension, body mass index (BMI) over 35 kg/m², age over 50 years, neck circumference over 40 cm, and male gender.¹³ Each question is answered with a "yes" or "no," with one point added for each "yes" response. A total score greater than 3 indicates a high risk for OSA and as the STOPBANG score increases, the probability of severe OSA increases.¹³

Modified Mallampati score

An airway assessment was conducted using the modified Mallampati method, where patients sat upright with their heads in a neutral position. They were asked to open their mouths widely and protrude their tongues fully. The airway was then classified based on the visible structures into four classes: Class I (soft palate, pillars, and tonsils were clearly visible), Class II (uvula, pillars, and upper poles of tonsils were visible), Class III (only part of the soft palate was visible; the tonsils, pillars, and base of the uvula could not be seen), and Class IV (only the hard palate was visible).¹⁴

Palatine tonsil grade

Tonsil size was graded from 0 to 4, where size 1 indicates tonsils hidden within the pillars, size 2 indicates tonsils extending to the pillars, size 3 indicates tonsils extending beyond the pillars but not reaching the midline, and size 4 indicates tonsils extending to the midline.¹⁵

Retrolingual and retropalatal Müller grading

Patients were examined in a sitting position using a flexible nasopharyngoscope to evaluate the degrees of retroglossal and retropalatal obstructions during the Müller maneuver, with computer-assisted area measurement calculating the extent of airway collapse. The cross-sectional area of the retroglossal level was measured both in a relaxed state (normal breathing, Ar) and in an active state (vigorous inhalation with closed mouth and nostrils occluded by an examiner, Aa). The retrolingual Müller grade and retropalatal Müller grade were then determined using the formula: $(Ar - Aa) / Ar \times 100\%$. The grading scale was as follows: 0 for a decrease in cross-sectional area of less than 25%, 1 for a decrease between

25% and 50%, 2 for a decrease between 50% and 75%, and 3 for a decrease of more than 75%.¹⁴

Statistical analysis

The statistical analysis was performed using SPSS version 26. All variables were analysed for their significance in relation to the Apnoea Hypopnea Index (AHI) through p-values ($p < 0.05$). Univariate and multivariate analyses were examined to explore the individual and combined effects of these variables on AHI value. Multicollinearity between statistically significance ($p < 0.05$) variables was checked, and for pairs with high multicollinearity (correlation > 0.7), only one variable from each pair was selected. Then, all variables that exhibited statistical significance were categorised into two sets of combined variables: Clinical Examination Score (CES) and ENT Examination Score (EES). The diagnostic accuracy of all screening tools (STOPBANG, ESS, CES, and EES), were calculated as the sensitivity and specificity. Receiver Operating Characteristic (ROC) curves were generated to analyse and compare the OSA diagnostic performance between STOPBANG, ESS, CES, and EES.

RESULTS

Participants' characteristics

A total of 201 patients were included in the study, with an average age of 47.8 ± 15.4 years, ranging from 16 to 89 years. Of these, 125 (62.2%) were male. The majority of the patients were Malay (116, 57.7%), followed by Chinese (51, 25.4%), Indian (32, 15.9%), and 2 (1.0%) from other races. The mean total AHI was 49.0 ± 33.5 , of which 127 (63.2%) were diagnosed with severe sleep apnoea ($\text{AHI} \geq 30$), whereas the mean total ESS was 10.3 ± 5.6 . The mean BMI was 33.9 ± 8.6 kg/m², neck circumference 40.9 ± 6.9 cm, and waist circumference 105.8 ± 20.8 cm. (Table 1)

Factors correlated with high AHI

Demographic factors correlated to high AHI included males (56.32 ± 30.4 vs 37.6 ± 35.2 in females, $p < 0.001$). Clinical factors included frequent snoring, the presence of tiredness, apnoea, nocturia and drooling. Anthropometric factors were weight ($r = 0.35$, $p < 0.001$), neck circumference ($r = 0.37$, $p < 0.001$), BMI ($r = 0.33$,

Table 1: Demographic and anthropometric characteristics of the cohort, N=201

Demographic		N (%)	Mean \pm SD
Age (years)			47.8 \pm 15.4
Gender	Female	76 (37.8)	
	Male	125 (62.2)	
Ethnic	Chinese	51 (21.4)	
	Indian	32 (15.9)	
	Malay	116 (57.7)	
	Others	2 (1.0)	
Anthropometric			
Weight (kg)			91 \pm 25.5
Height (cm)			164.3 \pm 9.8
BMI (kg/m ²)			33.9 \pm 8.6
Neck Circumference (cm)			40.9 \pm 6.9
Waist Circumference (cm)			105.8 \pm 20.8
Sleep			
Total ESS			10.3 \pm 5.6
Total AHI			49.2 \pm 33.5
Severity of sleep apnoea	Normal	9 (4.5)	
	Mild	22 (10.9)	
	Moderate	43 (21.4)	
	Severe	127 (63.2)	
Severe sleep apnoea	Yes ($\text{AHI} \geq 30$)	127 (63.2)	
	No ($\text{AHI} < 30$)	74 (36.8)	

p<0.001) and waist circumference (r=0.38, p<0.001). The Epworth Sleepiness Scale were Q1 (Question 1 of ESS: During sitting and reading) (r=0.17, p=0.014), Q3 (Question 3 of ESS: During Sitting inactive in public) (r=0.23, p<0.001), Q6 (Question 6 of ESS: During sitting and talking to someone) (r=0.17, p=0.015) and total ESS score

(r=0.20, p=0.005). ENT Assessment included Modified Mallampati (r=0.26, p=0.001), Palatine Tonsil grade (r=0.27, p<0.001), Retropalatal grade (r=0.29, p<0.001) and Retrolingual grade (r=0.29, p=0.002) were correlated with AHI. (Tables 2 and 3)

Table 2: Comparison means study between the categorical variables and AHI value

Demographic Characteristics		N	Mean ± SD	p-value
Age > 50 (years)	Yes	85	52.8±35.7	0.081
	No	116	44±29.7	
Gender	Female	76	37.6±35.2	<0.001
	Male	125	56.32±30.4	
Self-reported data				
Tiredness	Yes	112	53.8±35.8	0.044
	No	89	43.9±29.6	
Snore	Yes	188	50.4±34.0	0.098
	No	10	32.4±18.13	
Frequency of Snoring	Frequent	130	54.8±34.9	0.001
	No or rarely	32	32.0±27.2	
Stop breathing (apnoea)	Yes	118	52.9±34.2	0.032
	No	78	42.4±31.9	
Frequency of Apnoea	No or rarely	84	50.6±34.6	0.165
	Frequent	27	61.2±33.1	
Nocturia	Yes	155	52.8±33.4	0.002
	No	43	35.1±30.8	
Frequency of nocturia	1	59	44.4±30.6	0.053
	2	50	55.0±35.3	
	3	39	63.1±34.9	
	4	4	44.9±25.5	
Drooling of saliva	Yes	98	60.5±36.5	<0.001
	No	98	37.0±25.6	
Anthropometric Characteristics				
BMI > 35kg/m² (BMI35)	Yes	66	62.5±37.3	<0.001
	No	98	40.8±27.4	
Neck Circumference >40cm (NC40)	Yes	98	62.2±33.6	<0.001
	No	64	32.9±26.3	
ENT assessment				
Septum (Present of Deviated Nasal Septum (DNS)/Spur)	Yes	86	50.2±32.3	0.416
	No	32	44.8±30.9	
Inferior Turbinate	Abnormal	59	51.0±32.1	0.214
	Normal	82	44.6±28.9	
Polyps	Yes	1	70.0±0.0	0.528
	No	154	48.8±33.4	
Uvula	Abnormal	57	52.5±32.2	0.310
	Normal	121	47.0±34.3	

Table 3: Correlation between clinical variables and total AHI value

Variables	N	R	P-value
Demographic and Anthropometric			
Age	201	-0.064	0.366
Weight	156	0.35	<0.001
Height	157	0.11	0.157
BMI	151	0.33	<0.001
Neck circumference	162	0.37	<0.001
Waist circumference	141	0.38	<0.001
Total sleep duration	201	0.13	0.057
Epworth Sleepiness Scale (ESS)			
Q1 (Question 1 of ESS: During sitting and reading)	201	0.17	0.014
Q2 (Question 2 of ESS: During watching TV)	200	0.13	0.063
Q3 (Question 3 of ESS: During Sitting inactive in public)	200	0.23	0.001
Q4 (Question 4 of ESS: During as a passenger)	201	0.14	0.056
Q5 (Question 5 of ESS: During lying down to rest in the afternoon)	201	0.05	0.459
Q6 (Question 6 of ESS: During sitting and talking to someone)	201	0.17	0.015
Q7 (Question 7 of ESS: During sitting quietly after lunch)	200	0.11	0.115
Q8 (Question 8 of ESS: During in a car, while stopping at the traffic)	201	0.11	0.138
Total score of ESS	201	0.20	0.005
ENT assessment			
Modified Mallampati score	176	0.26	0.001
Friedman Tongue	179	0.13	0.084
Palatine Tonsil grade	173	0.27	<0.001
Lingual Tonsil grade	103	0.15	0.134
Retropalatal grade	165	0.29	<0.001
Retrolingual grade	161	0.29	0.002

Regression analysis

Regressing analysis

A hierarchical multiple regression analysis with 12 significant independent variables and one dependent variable (total AHI) was conducted across three models, as shown in Table 4. Model 1 included demographic and anthropometric variables (Gender, BMI >35kg/m² (BMI35), and Neck circumference >40cm (NC40). Waist circumference was excluded due to a strong correlation with weight ($r=0.81$). Neck circumference was incorporated as a separate variable because of its a modest correlation coefficient of 0.42 with BMI. Model 2 added self-reported data variables (Snore frequency, Stop breathing (Apnoea), Nocturia, Drooling saliva, and Tiredness). Model 3 incorporated ENT clinical examination variables (Palatine Tonsil grade, Modified Mallampati grade, Retropalatal grade, and Retrolingual grade).

The regression analysis, demographic and anthropometric variables (Step 1) explained 16.7%

of the variance in AHI. Adding self-reported data (Step 2) increased the variance explanation by 9.8% (total $R^2=26.5\%$, $F(8,192)=8.505$, $p<0.001$). Including ENT clinical examination variables (Step 3) further increased the variance by 6.5% (total $R^2=33.0\%$, $F(12,188)=7.20$, $p<0.01$). Significant changes in R^2 were noted in Steps 2 and 3 ($p=0.001$ and $p=0.005$). In the final model, significant variables were Palatine Tonsil grade ($\beta=0.239$, $p=0.001$), drooling saliva ($\beta=0.221$, $p=0.002$), BMI35 ($\beta=0.176$, $p=0.018$), and Gender ($\beta=0.172$, $p=0.019$).

Clinical Examination Score (CES) and ENT Examination Score (EES)

Two scores were formulated based on the regression analysis. The clinical examination score (CES) consisted of 8 variables: tiredness, apnoea, BMI >35, Neck circumference >40cm, Male, Snore frequency, Nocturia, and saliva drooling (Supplementary Figure 1). A score of

Table 4: Hierarchical regression analysis of predictors of total AHI values

Predictor	Regression					ANOVA	
	β	P	R ²	ΔR^2	P (ΔR^2)	F	P
Model 1							
Gender (male)	0.218	0.005	0.167	0.167	-	11.3	<0.001
BMI35	0.230	0.004					
NC40	0.144	0.089					
Model 2							
Gender (male)	0.173	0.022	0.265	0.098	0.001	7.35	<0.001
BMI35	0.211	0.006					
NC40	0.107	0.194					
Frequency of Snoring (frequent)	0.082	0.261					
Stop breathing (apnoea)	0.048	0.513					
Nocturia	0.106	0.126					
Drooling	0.231	0.002					
Tiredness	0.027	0.706					
Model 3							
Gender (male)	0.172	0.019	0.330	0.065	0.005	6.52	<0.001
BMI35	0.176	0.018					
NC40	0.066	0.416					
Frequency of Snoring (frequent)	0.062	0.393					
Apnoea	0.041	0.567					
Nocturia	0.101	0.139					
Drooling	0.221	0.002					
Tiredness	0.003	0.971					
Palatine Tonsil Score	0.239	0.001					
Modified Mallampati Score	0.051	0.473					
Retropalatal grade	0.028	0.719					
Retrolingual grade	0.030	0.692					

1 was given to each variable, with a maximum score of 8. The ENT examination score (EES) included 4 variables: palatine tonsil grade, Modified Mallampati score, retropalatal grade, and retrolingual grade (Supplementary Figure 2). Each variable was scored from 0 to 2 points, with a maximum score of 8. The correlation analysis showed that the CES model had the highest correlation coefficient ($R = 0.475$) and explained the most variance in total AHI ($R^2 = 22.6\%$), followed by STOPBANG ($R = 0.349$, $R^2 = 12.2\%$), ESS ($R = 0.198$, $R^2 = 3.9\%$), and EES ($R = 0.265$, $R^2 = 7.0\%$). (Table 5)

Sensitivity and specificity of CES, EES and STOPBANG

In the ROC curve analysis (Figure 1), the CES model had the highest area under the curve (AUC) at 0.735. This was followed by the STOPBANG model with an AUC of 0.709. The EES model had an AUC of 0.624, indicating moderate predictive performance. The ESS model showed the lowest discriminative ability with an AUC of 0.562.

Table 6 presents the sensitivity and specificity values for various cut-off scores of the Clinical Examination Score (CES) and compares them with

Table 5: Correlation analysis with total AHI

	N	Mean	Std. Deviation	R	R ²	p-value
Total AHI	201	49.23	33.49			
Clinical Examination Score (CES) ^a		4.40	1.90	0.475	0.226	<0.001
STOPBANG ^b		4.28	1.54	0.349	0.122	<0.001
ESS ^c		10.31	5.57	0.198	0.039	0.005
ENT Examination Score (EES) ^d		4.30	2.35	0.265	0.070	<0.001

^aCES (8 variables): Tiredness, Apnoea, BMI >35, Neck circumference >40cm, Male, Snore frequency, Nocturia and Drooling of saliva

^bSTOPBANG (8 variables): Snore, Tiredness, Apnoea, High Blood pressure, BMI >35, Age >50, Neck circumference >40cm and Male

^cESS : Epworth Sleepiness Scale, consists of 8 questions of different situation to fall asleep.

^dEES (4 variables): Modified Mallampati, Retropalatal grade, Retrolingual grade and Palatine tonsil grade.

the STOP-BANG score and ENT Examination Score (EES). At the lowest cut-off, the sensitivity is maximized but with lower specificity. As the cut-off score increases, sensitivity decreases while specificity improves, indicating a trade-off between identifying true positives and minimizing false positives. For CES, as the cut-off score increases, sensitivity drops while specificity improves. From scores 1 to 4, sensitivity decreases from 100% to 82.7%, and specificity increases from 5.4% to 50%. Starting at score 5, sensitivity sharply drops to 59.8%, and specificity rises to 74.3%. In contrast, the STOP-BANG score shows consistently high sensitivity, with values ranging from 100% at a score of ≥ 1 to 63.0% at ≥ 5 , but lower specificity. Specificity dramatically

increases to 71.6% at a score of ≥ 5 . For the EES, sensitivity ranges from 89.8% at a score of ≥ 1 to 59.8% at ≥ 5 , while specificity improves from 12.2% at ≥ 1 to 59.5% at ≥ 5 . Overall, CES offers very high specificity at every cut-off, while STOP-BANG offers high sensitivity, and EES provides moderate sensitivity and specificity.

Comparison among CES, STOPBANG, ESS and EES

Table 7 summarises the performance of different tools for detecting both moderate (AHI ≥ 15) and severe obstructive sleep apnoea (AHI ≥ 30). For detecting severe OSA (AHI ≥ 30), the STOP-BANG questionnaire shows high sensitivity

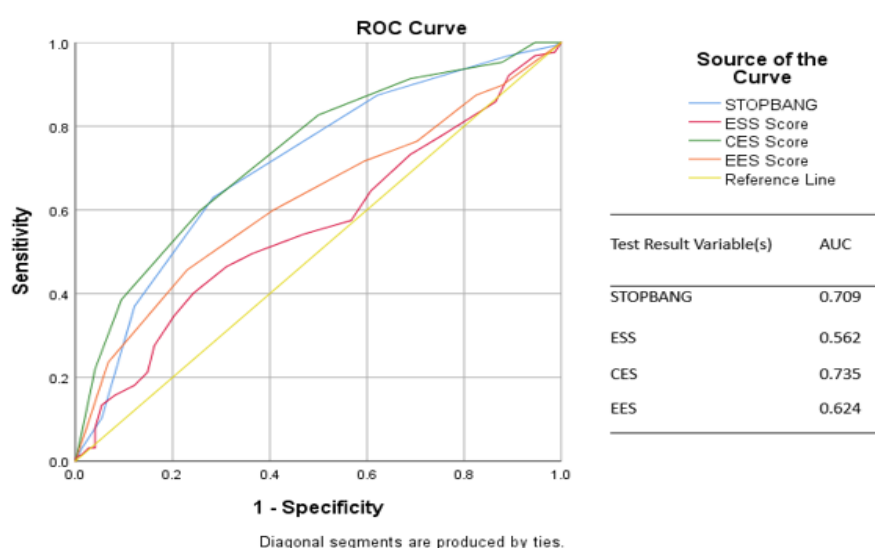


Figure 1: Area Under the Curve of ROC

Table 6: Sensitivity and specificity of different CES, STOPBANG and EES cutoffs

Score	Clinical Examination Score (CES)		STOPBANG		ENT Examination Score (EES)	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
≥ 1	100	5.4	100.0	0.0	89.8	12.2
≥ 2	95.3	12.2	99.2	1.4	87.4	17.6
≥ 3	91.3	31.1	96.9	10.8	76.4	29.7
≥ 4	82.7	50	87.4	37.8	71.7	40.5
≥ 5	59.8	74.3	63.0	71.6	59.8	59.5
≥ 6	38.6	90.5	37.0	87.8	45.7	77.0
≥ 7	22	95.9	10.2	94.6	23.6	93.2
= 8	7.1	98.6	3.1	98.6	89.8	12.2

(96.9%) but low specificity (10.8%). The Epworth Sleepiness Scale (ESS) has moderate specificity (63.5%) but low sensitivity (49.6%). The Clinical Examination Score (CES) strikes a good balance with high sensitivity (82.7%) and moderate specificity (50.0%). The ENT Examination Score (EES) shows moderate sensitivity (71.7%) but lower specificity (40.5%). For detecting moderate OSA (AHI ≥15), STOP-BANG has 95.9% sensitivity and 16.1% specificity, while CES shows 76.5% sensitivity and 61.3% specificity.

DISCUSSION

The study presented a novel approach to the assessment of sleep apnoea by introducing a Clinical Examination Score (CES) that incorporates various factors, including gender, BMI>35kg/m² (BMI35), Tiredness, neck circumference>40cm (NC40), observed apnoea, snore frequency, nocturia, and drooling. The results demonstrated that the CES, with a threshold of 4 or more, had a high sensitivity (82.7%) and moderate specificity (50%) in detecting severe sleep apnea, in contrast to the Epworth Sleepiness Scale (ESS), which had

lower sensitivity and moderate-high specificity.⁷ Choosing a cut-off score of 4 for the CES is justified based on the observed balance between sensitivity and specificity, with relatively high sensitivity, ensuring the detection of most true positive cases with significantly high specificity, thereby reducing the number of false positives.

Furthermore, the CES exhibited a significantly higher specificity than the STOPBANG questionnaire. Although widely used as a screening tool for sleep apnoea, STOPBANG has been noted to have low specificity¹⁶, leading to unnecessary testing and potentially overwhelming the PSG test facilities and a long waiting list. Based on the findings, the CES exhibits a specificity of 50.0% for AHI ≥30 and 61.3% for AHI ≥15, which is significantly higher compared to the STOPBANG questionnaire's specificity in this study (10.8% for AHI ≥30 and 16.1% for AHI ≥15). This higher specificity indicates that the CES is more adept at correctly identifying patients who have severe OSA and reducing the rate of false positives. This balance between sensitivity and specificity is crucial in clinical settings with limited resources

Table 7: Sensitivity and specificity of different groups of variables to detect moderate AHI (AHI≥ 15) and Severe AHI (≥30)

Variables	AHI ≥ 15		AHI ≥ 30	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
STOPBANG	95.9	16.1	96.9	10.8
ESS	54.7	71.0	49.6	63.5
Clinical Examination Score (CES)	76.5	61.3	82.7	50.0
ENT Examination Score (EES)	69.4	45.2	71.7	40.5

and a long waiting list. Furthermore, it ensures that patients more likely to have severe OSA are prioritised for early testing, enhancing the overall efficiency of sleep disorder management.

The Clinical Examination Score (CES) has some common variables with the STOPBANG questionnaires, including Tiredness (T), Observed apnoea (O), Neck circumference (N), BMI (B), and Gender (G).¹³ Age has been considered a major risk factor for sleep apnoea, with its prevalence increasing as individuals get older.¹⁷ However, in this study, age (A in STOPBANG) was not significantly correlated with AHI, likely because the patients were diagnosed with sleep apnoea at a younger age (< 50 years). We have replaced the presence of snoring (S) with the frequency of snoring, which is more specific and occasional snoring may result from temporary factors like alcohol consumption, sleeping position, or nasal congestion.¹⁸ In contrast, frequent snoring suggests a persistent obstruction of the upper airway, a hallmark of sleep apnoea.¹⁹ Blood pressure (P) is not statistically correlated with AHI in our study (Supplementary Table 1), likely because it is prevalent in our population.²⁰

The relationship between the “Tiredness” component in the STOPBANG questionnaire and the Epworth Sleepiness Scale (ESS) is notable because both assess symptoms commonly associated with sleep apnea, such as daytime sleepiness and fatigue.⁷ The “Tiredness” component of STOPBANG directly inquires if an individual often feels tired, fatigued, or sleepy during the day, addressing a primary symptom of sleep apnea.²¹ On the other hand, the ESS offers a more detailed evaluation of daytime sleepiness across different situations, providing a broader perspective on sleep-related fatigue.¹² Despite the comprehensive nature of the ESS, its lower sensitivity limits its utility in detecting severe sleep apnea.⁷ In this study, only three out of the eight ESS questions were statistically significant: Q1 (while sitting and reading), Q3 (sitting inactive in public), and Q6 (sitting and talking to someone). This suggests that these specific situations are more indicative of the tiredness associated with sleep apnea in Malaysian context. Therefore, combining these significant ESS questions into a single variable (Tiredness) could improve the accuracy of sleep apnea screening by capturing both general and situational tiredness.

In this study, drooling of saliva and nocturia were found to be significantly correlated with higher AHI. Drooling during sleep is one of the symptoms commonly associated with sleep

apnoea^{22,23}, because of mouth breathing²⁴ and reduced swallowing during sleep.²⁵ Nocturia is closely related to arousals during apnoeic episodes.²⁶ Furthermore, negative thoracic pressure during obstructed breathing can lead to increased release of atrial natriuretic peptide (ANP) and decreased secretion of arginine vasopressin (AVP), resulting in increased urine production.²⁷

The ENT Examination Score (EES) offers valuable insights into detecting upper airway structural obstructions that may exacerbate sleep apnea. In this study, the EES was created with balances sensitivity and specificity in detecting obstructive sleep apnea (OSA). With a cut-off score of 4, it has 71.7% sensitivity and moderate 40.5% specificity. As the EES threshold score increases, the specificity improves significantly, showing better accuracy in excluding patients who do not have OSA. Higher thresholds improve specificity significantly; for instance, a score of 6 has 77% specificity, and 7 reaches 93.2%, although sensitivity drops. This trade-off means lower scores are better for broadly identifying potential OSA cases, while higher scores are more precise in ruling out those without OSA.

The EES incorporates components such as palatine tonsil grade, Modified Mallampati score, and retropalatal and retrolingual grades, all of which have been found to have a statistically significant correlation with AHI values, which is compatible with previous studies.^{28,29} The risk of obstruction depends on these structures' airway adherence, the ability of the airway to expand and contract in response to changes in pressure. The palatine tonsil grade is a clinical measure used to evaluate the size of the tonsils. Enlarged tonsils, especially in Grades 3 and 4, reduce the diameter of the oropharyngeal airway, can cause partial or complete obstruction during sleep, leading to apneas and hypopneas.³² Friedman, *et al.*²⁹ mention in their study, that tonsil size was significantly predictive of OSA and could contribute to gauging the severity of the condition and as a commonly targeted surgical site in the treatment of OSA. The Mallampati score is a straightforward and widely utilized clinical tool for assessing the anatomical features of a patient's airway, predicting both the ease of intubation and the severity of OSA.²⁹ Numerous studies, including the present one, have demonstrated a significant correlation between higher Mallampati scores, especially grade 3 and 4, with increased Apnea-Hypopnea Index (AHI) values. Reduced adherence in the retropalatal and retroglottal

regions can lead to increased collapsibility of the airway walls, especially during sleep when muscle tone decreases. Higher grades in all these ENT assessments are associated with increased apnea-hypopnea index, signifying the significance of their role in contributing to OSA severity and serving as a strong predictor of OSA severity.

For centres with specialized ENT expertise, incorporating the EES as a diagnostic tool could enhance OSA evaluations by providing critical structural information that complements symptom-focused assessments like the STOP-BANG, Epworth Sleepiness Scale (ESS) or the new Clinical Examination Score (CES). For example, by implementing a serial diagnostic approach, from CES (cut off ≥ 4) to EES (cut off ≥ 4), overall specificity will increase to 72.5%, calculated using the formula for serial testing specificity — specificity: $\frac{A_{CES}}{A_{CES} + B_{EES}}$ (given: A_{CES} specificity = 50.0% and B_{EES} specificity = 40.5%).³⁰ This enhanced specificity allows patients who screen positive to proceed directly to titration studies for treatment, bypassing the need for a Level 1 PSG for diagnostic confirmation. Instead, Level 2 or 3 PSG, which are less labor-intensive and more readily available, can suffice for confirmation. This streamlined pathway significantly reduces the burden on Level 1 PSG resources and facilitates faster diagnosis and treatment for patients requiring immediate care.

A limitation of this study is its retrospective design, relying on previously collected data for analysis. Since the data was gathered in the past, it may not fully reflect any new clinical variables or predictive markers identified. Therefore, validation for CES and EES in future studies would be necessary to confirm their effectiveness in real-time patient assessments. This approach ensures that any advancements are accurately evaluated in current clinical settings, offering reliable results for practical use.

In conclusion, integrating the Clinical Examination Score (CES) and the ENT Examination Score (EES) significantly enhances current screening and management of moderate to severe sleep apnea. Improved specificity helps reduce false positives, which in turn alleviates the burden on PSG resources and shortens waiting times. This approach also prioritizes individuals with severe cases, promoting more timely diagnosis and treatment, which can ultimately improve outcomes in managing sleep apnea syndrome (SAS).

ACKNOWLEDGMENTS

The authors wish to acknowledge the invaluable support of the staff at the UMMC Sleep Laboratory and the medical officers at the UMMC ENT Clinic for their assistance in data provision.

DISCLOSURE

Data availability: The datasets generated and analysed during the current study are not publicly available due to patient confidentiality but are available from the corresponding author upon reasonable request. Access to the data will be granted in accordance with applicable legal and ethical guidelines, ensuring the protection of patient privacy.

Financial support: None

Conflict of interest: None

REFERENCES

1. Arens R, Marcus CL. Pathophysiology of upper airway obstruction: a developmental perspective. *Sleep* 2004;27(5): 997-1019. doi: 10.1093/sleep/27.5.997.
2. Hirsch Allen AJM, Bansback N, Ayas NT. The effect of OSA on work disability and work-related injuries. *Chest* 2015;47(5):1422-8. doi: 10.1378/chest.14-1949.
3. Marin-Oto M, Vicente EE, Marin JM. Long term management of obstructive sleep apnea and its comorbidities. *Multidiscip Respir Med* 2019;14(1):21. doi: 10.1186/s40248-019-0186-3.
4. Jean-Louis G, Zizi F, Clark LT, Brown CD, McFarlane SI. Obstructive sleep apnea and cardiovascular disease: role of the metabolic syndrome and its components. *J Clin Sleep Med* 2008; 4(3):261-72.
5. Srijithesh PR, Aghoram R, Goel A, Dhanya J. Positional therapy for obstructive sleep apnoea. *Cochrane Database Syst Rev* 2019; 5(5):CD010990. doi: 10.1002/14651858.cd010990.pub2.
6. Yeghiazarians Y, Jneid H, Tietjens JR, et al. Obstructive sleep apnea and cardiovascular disease: A scientific statement from the American Heart Association. *Circulation* 2021;144(3):e56-e67. doi: 10.1161/cir.0000000000000988.
7. Zheng ZZ, Zhang YT, Chen MD, et al. Application value of joint STOP-Bang questionnaire and Epworth Sleepiness Scale in screening for obstructive sleep apnea. *Front Public Health*. 2022;10:950585. doi: 10.3389/fpubh.2022.950585.
8. Loh JMR, Toh ST. Rethinking neck circumference in STOP-BANG for Asian OSA. *Proceedings of Singapore Healthcare* 2018; 28(2):105-9. doi: 10.1177/2010105818810272.
9. Bonzelaar LB, Salapatras AM, Yang J, Friedman M. Validity of the epworth sleepiness scale as a screening tool for obstructive sleep apnea. *Laryngoscope* 2017;127(2):525-31. doi: 10.1002/lary.26206.

10. Berry RB, Brooks R, Gamaldo C, *et al.* AASM Scoring Manual Updates for 2017 (Version 2.4). *J Clin Sleep Med* 2017;13(5):665-6. doi: 10.5664/jcsm.6576.
11. Budhiraja R, Javaheri S, Parthasarathy S, Berry RB, Quan SF. The association between obstructive sleep apnea characterized by a minimum 3 percent oxygen desaturation or arousal hypopnea definition and hypertension. *J Clin Sleep Med* 2019;15(9):1261-70. doi: 10.5664/jcsm.7916.
12. Smith SS, Oei TP, Douglas JA, Brown I, Jorgensen G, Andrews J. Confirmatory factor analysis of the Epworth Sleepiness Scale (ESS) in patients with obstructive sleep apnoea. *Sleep Med* 2008;9(7):739-44. doi: 10.1016/j.sleep.2007.08.004.
13. Chung F, Abdullah HR, Liao P. STOP-Bang questionnaire: A practical approach to screen for obstructive sleep apnea. *Chest* 2016;149(3):631-8. doi: 10.1378/chest.15-0903.
14. Wu MJ, Ho CY, Tsai HH, Huang HM, Lee PL, Tan CT. Retropalatal Müller grade is associated with the severity of obstructive sleep apnea in non-obese Asian patients. Retropalatal Müller grade and OSA in non-obese. *Sleep Breath* 2011;15(4):799-807. doi: 10.1007/s11325-010-0441-5.
15. Friedman M, Ibrahim H, Bass L. Clinical staging for sleep-disordered breathing. *Otolaryngol Head Neck Surg* 2002;127(1):13-21. doi: 10.1067/mhn.2002.126477.
16. 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 Netw Open* 2021;4(3):e211009. doi: 10.1001/jamanetworkopen.2021.1009.
17. Ralls FM, Grigg-Damberger M. Roles of gender, age, race/ethnicity, and residential socioeconomic in obstructive sleep apnea syndromes. *Curr Opin Pulm Med* 2012; 18(6):568-73. doi: 10.1097/MCP.0b013e328358be05.
18. Osman EZ, Osborne J, Hill PD, Lee BWV. The Epworth Sleepiness Scale: can it be used for sleep apnoea screening among snorers? *Clin Otolaryngol Allied Sci* 1999;24(3):239-41. doi: 10.1046/j.1365-2273.1999.00256.x.
19. Maimon N, Hanly PJ. Does snoring intensity correlate with the severity of obstructive sleep apnea? *J Clin Sleep Med* 2010;6(5):475-8. doi: 10.5664/jcsm.27938.
20. Abougambou SSI, Abougambou AS, Sulaiman SAS, Hassali MA. Prevalence of hypertension, control of blood pressure and treatment in hypertensive with type 2 diabetes in Hospital University Sains Malaysia. *Diabetes Metab Syndr* 2011;5(3): 115-9. doi: 10.1016/j.dsx.2012.03.001.
21. Chung F, Yeneswaran B, Liao P, *et al.* STOP Questionnaire: A tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008; 108(5):812-21. doi: 10.1097/aln.0b013e31816d83e4.
22. Kato T. Sleep bruxism and its relation to obstructive sleep apnea-hypopnea syndrome. *Sleep and Biological Rhythms* 2004;2(1):1-15. doi: https://doi.org/10.1111/j.1479-8425.2003.00077.x.
23. Arali V, Namineni S, Sampath C. Pediatric obstructive sleep apnea syndrome: Time to wake up. *Int J Clin Pediatr Dent* 2012;5(1):54-60. doi: 10.5005/jp-journals-10005-1134.
24. Tsuda H, Lowe AA, Chen H, Fleetham JA, Ayas NT, Almeida FR. The relationship between mouth opening and sleep stage-related sleep disordered breathing. *J Clin Sleep Med* 2011;7(2):181-6. doi: 10.5664/jcsm.28107.
25. Lobbezoo F, Aarab G, Wetselaar P, Hoekema A, de Lange J, de Vries N. A new definition of dental sleep medicine. *J Oral Rehabil* 2016;43(10):786-90. doi: 10.1111/joor.12421.
26. Pressman MR, Figueroa WG, Kendrick-Mohamed J, Greenspon LW, Peterson DD. Nocturia: A rarely recognized symptom of sleep apnea and other occult sleep disorders. *Arch Int Med* 1996; 156(5):545-50. doi: 10.1001/archinte.1996.00440050103011.
27. Umlauf MG, Chasens ER, Greevy RA, Arnold J, Burgio KL, Pillion DJ. Obstructive sleep apnea, nocturia and polyuria in older adults. *Sleep* 2004;27(1):139-44. doi: 10.1093/sleep/27.1.139.
28. Brennan L, Kirkham FJ, Gavlak JC. Sleep-disordered breathing and comorbidities: role of the upper airway and craniofacial skeleton. *Nat Sci Sleep* 2020; 12:907-36. doi: 10.2147/nss.s146608.
29. Friedman M, Tanyeri H, La Rosa M, *et al.* Clinical predictors of obstructive sleep apnea. *Laryngoscope* 1999;109(12):1901-7. doi: https://doi.org/10.1097/00005537-199912000-00002.
30. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. *AJR Am J Roentgenol* 2005;184(1):14-9. doi: 10.2214/ajr.184.1.01840014.

Clinical Examination Score (CES)
(Sleep Apnea risk assessment)

		Score: 0		Score: 1	
1.	Gender	Female		Male	
2.	BMI	< 35 kg/m ²		≥35 kg/m ²	
3.	Neck Circumference	< 40cm		≥ 40cm	
4.	Snore frequency	No/rarely		Frequent/very frequent	
5.	Do you feel Tired or sleepy during sitting, reading, or during talking with someone? (Daytime)	No		Yes	
6.	Do you stop breathing in sleep?	No		Yes	
7.	Do you wake up at night to go to the toilet ?	No		Yes	
8.	Do you have drooling while you sleep? (Pillow wet by saliva)	No		Yes	

Total score = ____/8

ENT Examination Score (EES) (Sleep Apnea risk assessment)							
		Score: 0		Score: 1		Score: 2	
1.	Mallampati Grade	G1		G2		G3/ 4	
2.	Palatine Tonsil	G1		G2		G3/ 4	
3.	Retropalatal Grade	G0		G1		G2/ 3	
4.	<u>Retrolingual</u> Grade	G0		G1		G2/ 3	

Total score = ____/8

1) Mallampati Grade

- Grade 1: The soft palate, pillars, and tonsils were clearly visible
- Grade 2: The uvula, pillars, and upper poles of tonsils were visible.
- Grade 3: Only part of the soft palate was visible; the tonsils, pillars, and base of the uvula could not be seen.
- Grade 4: Only the hard palate was visible.

2) Palatine Tonsil Grade

- Grade 1: The tonsils are within the tonsillar fossa and do not extend past the pillars.
- Grade 2: The tonsils extend slightly beyond the pillars
- Grade 3: The tonsils extend beyond the pillars but not to the midline
- Grade 4: The tonsils extend to the midline

3) Retropalatal and Retrolingual Muller grade:

- Grade 0: Diminishing cross-sectional less than 25%
- Grade 1: Diminishing cross-sectional between 25% to 50%
- Grade 2: Diminishing cross-sectional between 50% to 75%
- Grade 3: Diminishing cross-sectional more than 75%

Figure S2: ENT Examination Score (EES)

Demographic Characteristics		N	Mean \pm SD	p-value
Ethnic	Chinese	51	42.1 \pm 27.1	0.432
	Indian	32	48.3 \pm 30.3	
	Malay	116	52.2 \pm 36.5	
	Others	2	48.3 \pm 61.5	
Sleep category	Early (sleep below 12pm)	169	49.2 \pm 33.5	0.893
	Late (sleep after 12pm)	27	50.2 \pm 35.2	
Presence of sleeping partner	Yes	161	49.0 \pm 33.4	0.835
	No	39	50.2 \pm 34.4	
Self-reported data				
Gaspings or choking	Yes	122	51.2 \pm 45.6	0.269
	No	72	45.6 \pm 33.5	
Frequency of gasping/choking	Frequent	21	59.6 \pm 37.0	0.272
	No or rarely	92	50.5 \pm 33.6	
Dry mouth	Yes	136	51.5 \pm 32.5	0.173
	No	62	44.4 \pm 35.5	
Heart Disease	Yes	33	45.7 \pm 30.4	0.544
	No	162	49.6 \pm 34.3	
High blood pressure	Yes	93	49.4 \pm 26.6	0.868
	No	102	48.5 \pm 39.2	
Lung disease	Yes	23	48.8 \pm 49.5	0.999
	No	172	48.9 \pm 31.2	
Kidney/Liver/GI disease	Yes	21	50.3 \pm 30.4	0.845
	No	173	48.8 \pm 34.2	
Endocrine Disease	Yes	49	49.1 \pm 37.1	0.999
	No	145	49.1 \pm 32.6	
Neurological Problem	Yes	17	48.7 \pm 28.9	0.968
	No	177	49.1 \pm 34.2	