Original Article

STOP-BANG Score versus Epworth Sleepiness Scale as a Screening Tool for Obstructive Sleep Apnea

Abstract

Background: Obstructive sleep apnea (OSA) is a common sleep-breathing disorder. OSA is becoming highly prevalent, which means that to detect and prevent various hazardous complications, it is imperative that there are easy yet accurate techniques available to identify people with OSA. **Objectives:** The objective was to compare two questionnaires: The STOP-BANG score and the Epworth Sleepiness Scale (ESS), used to screen the patients suspected of OSA. Polysomnography (PSG), the gold standard investigation, was used to diagnose OSA. **Materials and Methods:** Sixty-five suspected individuals were recruited as per inclusion/exclusion criteria. Detailed history taking, physical examination, and anthropometric examination were done in all patients. Suspects were subjected to filling up of the STOP-BANG and ESS questionnaires, following which they underwent an overnight PSG examination, which is considered the gold standard diagnostic investigation for OSA. Compiled data were used to compare the sensitivities, specificities, and positive and negative predictive values (NPVs) of the two screening scores. **Results:** Of 65 screened patients, 57 (88%) had OSA. The sensitivity to predict OSA was the highest for the STOP-BANG questionnaire (91.23%), whereas ESS had a sensitivity of 70.18%. No difference in specificity (75%) of the two scores was noted. The positive predictive values of STOP-BANG and ESS questionnaires was 96.30% and 95.20%, respectively. NPV of STOP-BANG and ESS was 54.50% and 26.10%, respectively. **Conclusion:** The present study was able to provide valuable insights into OSA screening. Out of the two studied OSA screening questionnaires, we found out that both had comparatively good predictive and diagnostic accuracy, with the STOP-BANG score surpassing the ESS score in the majority of measures. Considering the high global burden of undiagnosed OSA, there is a need to upregulate the screening for OSA followed by appropriate treatment measures. This would improve sleep quality and reduce the risk of complications and future adverse health outcomes.

Keywords: *Epworth Sleepiness Scale questionnaire, obstructive sleep apnea, screening, STOP-BANG*

Manpriya Dogra, Surabhi Jaggi[1], Komaldeep Kaur, Chahat Bhatia, Deepak Aggarwal, Varinder Saini

Department of Pulmonary, Critical Care and Sleep Medicine, Government Medical College and Hospital, Chandigarh, [1]Department of Pulmonary Medicine, Indira Gandhi Medical College, Shimla, Himachal Pradesh, India

Introduction

Obstructive sleep apnea (OSA) is a syndrome marked by apneas, hypopneas, and respiratory effort-related awakening from sleep. It involves repeated obstruction of the pharyngeal airway while asleep, with resultant hypoxemia and sleep fragmentation.

Based on apnea–hypopnea index (AHI), the American Academy of Sleep Medicine (AASM) defines OSA syndrome as an AHI >5 with associated symptoms like abnormally high levels of daytime sleepiness, fatigue, or impaired cognition or an AHI of >15, regardless of associated symptoms.[1]

The prevalence of OSA appears to be rising, which may be related to rising obesity rates or higher OSA identification rates. In numerous studies that have estimated sex-specific prevalence for OSA, a 2–3-fold enhanced risk for men compared to women has been reported.[2]

An increased incidence of dysrhythmias, pulmonary hypertension, biventricular failure, myocardial infarction, diurnal hypertension, stroke, and metabolic syndrome have been noted among individuals suffering from OSA. There has also been a higher incidence of cognitive impairment, which negatively affects work performance. This demands a timely detection and treatment of OSA to prevent detrimental effects on health.

How to cite this article: Dogra M, Jaggi S, Kaur K, Bhatia C, Aggarwal D, Saini V. STOP-BANG score versus Epworth Sleepiness Scale as a screening tool for obstructive sleep apnea. Int J App Basic Med Res 2025;15:116-21.

Submitted: 28-Nov-2024
Revised: 05-Feb-2025
Accepted: 11-Mar-2025
Published: 07-Apr-2025

Address for correspondence:
Dr. Surabhi Jaggi, Department of Pulmonary Medicine, Second Floor, D Block, Indira Gandhi Medical College, Shimla - 171 001, Himachal Pradesh, India. E-mail: surabhijaggi@gmail.com

Access this article online

Website: https://journals.lww.com/IJAB

DOI: 10.4103/ijabmr.ijabmr_557_24

Quick Response Code:

The gold standard diagnostic test for sleep disorders, including OSA, is polysomnography (PSG). It is a systematic process of gathering physiologic variables, while the patient sleeps attached to various monitoring devices. From these physiologic variables, an abundance of information is derived. This includes sleep architecture, the frequency of abnormal events (e.g., apneas) during sleep, and various diagnostic measures (e.g., AHI).[1]

PSG is expensive and not available everywhere. With the help of common signs and symptoms, various prediction questionnaires, scores, and clinical prediction rules have been evaluated as screening tools for OSA. These can be obtained easily and elucidated at the primary care level.

The Epworth Sleepiness Scale (ESS) was designed by Dr. Johns in 1990 and subsequently modified slightly in 1997. The ESS is a one-page questionnaire that asks respondents to allocate one point each in eight different sedentary situations on the estimates of their likelihood of falling asleep or dozing off.[3] As OSA may affect airway management during surgery, the STOP-BANG questionnaire (SBQ) was designed in 2008 as a preoperative screening tool for anesthetists to check for undiagnosed OSA.[4] It is a questionnaire comprising eight items that incorporates the following information: presence of snoring, fatigue or tiredness, history of observed apneas, raised blood pressure, increase in body mass index (BMI), age, neck circumference, and gender. While filling out the questionnaire, one point is allotted to each symptom or risk factor, with a maximum score equaling to eight. A score between 0 and 2 confers a low risk for moderate-to-severe OSA, whereas a score between 5 and 8 is considered to be at high risk for this condition.[5]

In literature, evidence of screening abilities of various scores is documented, but there is a scarcity of comparative data on the Indian population. The purpose of this study was to evaluate and compare the sensitivity, specificity, and accuracy of the ESS and the SBQ. PSG was used to confirm the diagnosis of OSA for the same.

Materials and Methods

This was a cross-sectional study conducted in the Department of Pulmonary, Critical Care, and Sleep Medicine at Government Medical College and Hospital (GMCH), Chandigarh, over a period of 1½ years from December 2022 to May 2024. Sixty-five suspected patients of OSA more than 18 years of age, attending the pulmonary outpatient department (OPD), were consecutively enrolled in the study after taking informed written consent. Suspicion was based on the presence of excessive snoring and daytime sleepiness among the patients. Patients with other comorbid conditions such as chronic obstructive pulmonary disease, interstitial lung disease, lung cancer, history of infection in the last 4 weeks, and history of recent cardiovascular decompensation due to myocardial infarction or cardiac failure in the preceding 4 weeks were excluded from the study. The study was conducted after approval from the institute's ethics committee. Each subject, after a detailed medical history, was subjected to general physical examination including weight, height, waist, and neck circumference measurement and systemic examination. Thereafter, they were subjected to fill two OSA screening questionnaires. SBQ and ESS and their scores were recorded. After initial screening, subjects underwent an overnight PSG in the sleep laboratory under the Department of Pulmonary, Critical Care and Sleep Medicine, GMCH, using the Compumedics E-series-44 channel PSG system. In this diagnostic investigation, various parameters were recorded such as brain electrical activity (electroencephalogram), eye movements (oculogram), jaw muscle movements (electromyography), leg muscle movement, airflow, respiratory efforts, electrocardiogram, and oxygen saturation (SpO_2).

A diagnosis of OSA was made as per the guidelines of AASM which accepts the criterion of AHI > 05 on PSG.[6] The AHI is the total number of apneas/hypopneas per hour of sleep and an apneic event being defined as a cessation of oronasal airflow for at least 10 s, while hypopnea as a 50% decrease in oronasal airflow, associated with an oxyhemoglobin desaturation of more than 3%. OSA was categorized as mild, moderate, and severe OSA based on the AHI. Mild OSA is defined as AHI 5–15, moderate OSA as AHI 15–30, and severe OSA as AHI > 30.[7]

Compiled data were used to compare the sensitivities and specificities of the two screening scores, keeping the PSG as the gold standard.

Results

The demographic details of the subjects are mentioned in Table 1. The mean age of the study participants was 52.95 ± 11.60 years. Majority of patients belonged to middle and old-age groups. A male preponderance (n = 51, 78%) was observed among the study participants. The mean BMI of study participants was 30.54 ± 6.11 kg/m^2.

Table 1: Demographic details

Parameter	Group	Frequency, *n* (%)
Age (years)	Young (18–40)	11 (16.92)
	Middle (41–60)	38 (58.46)
	Older (>60)	16 (24.61)
Gender	Male	51 (78)
	Female	14 (22)
Addictions	Smoking	3 (4.62)
	Alcoholism	11 (16.92)
Comorbidities	Diabetes	15 (23.08)
	Hypertension	49 (75.38)
	Coronary artery disease	5 (7.69)
	Hypothyroidism	8 (12.31)

Hypertension (n = 49) was the most common comorbidity, followed by diabetes (n = 15) [Table 1].

In the study, the majority of cases (n = 46, 70.77%) were diagnosed with severe OSA. A smaller proportion (n = 9, 13.85%) had moderate OSA. Notably, 12.31% (n = 8) had a normal PSG finding, while only 3.08% (n = 2) were classified as having mild OSA. The mean AHI, which quantifies the severity of OSA, was 44.86, with a standard deviation of 25.97. The median AHI was reported as 42, with an interquartile range of 28–62.7 [Table 2].

The study recorded the mean STOP-BANG score as 5.82 with a standard deviation of 0.9. The median STOP-BANG score among the subjects was consistently reported at 6, with a narrow interquartile range precisely fixed at 6. In the study, the risk of OSA, as assessed by the SBQ was predominantly high (n = 54, 83.08%). Conversely, a smaller group (n = 11, 16.92%) was reported to have a low risk of OSA.

In the study, the proportion of patients classified according to the risk of OSA through the SBQ showed significant differences across various OSA severity groups. For the low-risk category, 75% of patients with no OSA were assessed as low risk, which was significantly higher compared to those with mild OSA (50%), moderate OSA (33.33%), and severe OSA (2.17%). Conversely, the proportion of patients assessed as high risk was markedly higher in the severe OSA group at 97.83%, compared to those with no OSA (25%), mild OSA (50%), and moderate OSA (66.67%). The differences observed were statistically significant, with $P < 0.0001$ [Table 3].

The area under the receiver operating characteristic (ROC) curve (AUC) of SBQ for predicting OSA was 0.781, with a standard error of 0.119 and a 95% confidence interval ranging from 0.547 to 1 [Figure 1]. The P value for this analysis was 0.0185, indicating statistical significance. The optimal cutoff point was identified as >5. At this cutoff, the sensitivity was 91.23% with a confidence interval of 80.7%–97.1%, and the specificity was 75% with a confidence interval of 34.9%–96.8%. The positive predictive value (PPV) was 96.3% (87.3%–99.5%), and the negative predictive value (NPV) was 54.5% (23.4%–83.3%). The overall diagnostic accuracy was 89.23%, with P = 0.016 from the Chi-square test [Table 4].

The study found the average ESS score among subjects to be 11.38, with a standard deviation of 2.98. The median ESS score was 11, with a 25th to 75th percentile range from 10 to 12. ESS scores indicated varying levels of daytime sleepiness among participants. The largest group, consisting of 35 cases or 53.85%, experienced mild sleepiness.

The AUC of ESS for predicting OSA was 0.729, with a standard error of 0.0718 and a 95% confidence interval from 0.589 to 0.870 [Figure 2]. The P value was 0.0014, also indicating statistical significance. The optimal cutoff was identified as >10. At this cutoff, the sensitivity was

Table 2: Apnea–hypopnea index distribution in study participants

AHI	Frequency (%)
No OSA	8 (12)
Mild OSA	2 (3.08)
Moderate OSA	9 (13.85)
Severe OSA	46 (70.77)
Mean±SD	44.86±25.97
Median (25th–75th percentile)	42 (28–62.7)
Range	3.1–105.2

OSA: Obstructive sleep apnea; SD: Standard deviation; AHI: Apnea–hypopnea index

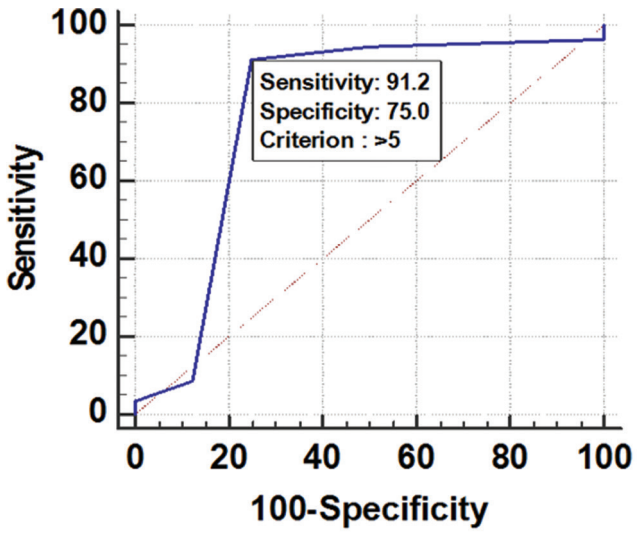

Figure 1: Receiver operating characteristic curve of STOP-BANG score for predicting obstructive sleep apnea

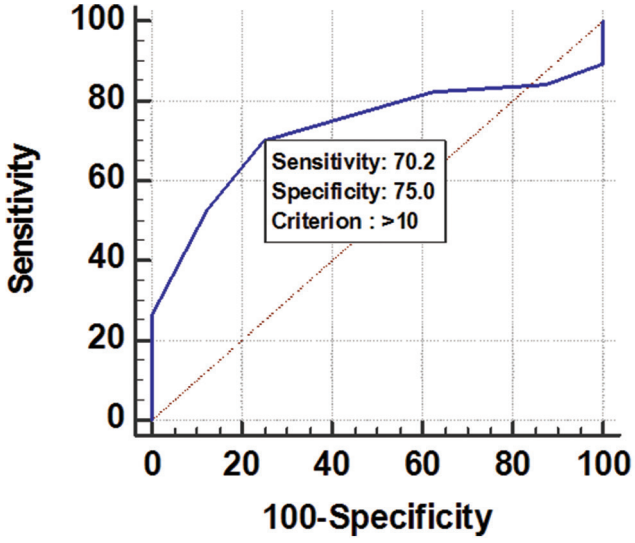

Figure 2: Receiver operating characteristic curve of Epworth Sleepiness Scale score for predicting obstructive sleep apnea. ESS: Epworth Sleepiness Scale

Table 3: Association of risk of obstructive sleep apnea according to STOP-BANG with severity of apnea–hypopnea index

Risk of OSA according to STOP-BANG	No OSA (*n*=8), *n* (%)	Mild OSA (*n*=2), *n* (%)	Moderate OSA (*n*=9), *n* (%)	Severe OSA (*n*=46), *n* (%)	Total	*P*
Low	6 (75)	1 (50)	3 (33.33)	1 (2.17)	11 (16.92)	<0.0001*
High	2 (25)	1 (50)	6 (66.67)	45 (97.83)	54 (83.08)	
Total	8 (100)	2 (100)	9 (100)	46 (100)	65 (100)	

**P* value less than 0.05 is considered significant. OSA: Obstructive sleep apnea

Table 4: Receiver operating characteristic curve of STOP-BANG score and Epworth Sleepiness Scale score for predicting obstructive sleep apnea

Variables	STOP-BANG score		ESS score
AUC	0.781		0.729
SE	0.119		0.0718
95% CI	0.547–1.000		0.589–0.870
P	0.0185		0.0014
Cutoff	>5		>10
Sensitivity (95% CI)	91.23% (80.7%–97.1%)		70.18% (56.6%–81.6%)
Specificity (95% CI)	75% (34.9%–96.8%)		75% (34.9%–96.8%)
PPV (95% CI)	96.3% (87.3%–99.5%)		95.2% (83.8%–99.4%)
NPV (95% CI)	54.5% (23.4%–83.3%)		26.1% (10.2%–48.4%)
Diagnostic accuracy	89.23%		70.77%
P value of diagnostic accuracy		0.016§	
P value of sensitivity		0.017‡	
P value of specificity		1‡	

‡McNemar test; §Chi-square test. SE: Standard error; CI: Confidence interval; NPV: Negative predictive value; PPV: Positive predictive value; AUC: Area under the ROC curve; ESS: Epworth Sleepiness Scale

70.18% with a confidence interval of 56.6%–81.6% and the specificity of 75% with the same confidence interval. The PPV was slightly lower at 95.2% (83.8%–99.4%), and the NPV was significantly lower at 26.1% (10.2%–48.4%). The overall diagnostic accuracy was 70.77%, with the *P* value for sensitivity being 0.017 from the McNemar test [Table 4].

Discussion

The study was conducted with the aim of comparing the prediction probabilities for OSA of two well-known sleep questionnaires: STOP-BANG score and ESS.

According to the PSG findings, mild OSA was reported in 3% (*n* = 2), moderate OSA in 14% (*n* = 9), and severe OSA in 71% (*n* = 46) patients, respectively. Conversely, 12% (*n* = 8) cases showed no signs of OSA. AHI ranged from 3.1 to 105.2 with a mean of 44.86 ± 25.97. A higher prevalence of severe OSA (70%) was seen in our study as we enrolled patients from OPD with clinical suspicion of OSA, presenting with symptoms such as snoring and excessive daytime sleepiness. As per literature, the occurrence of OSA in snoring patients ranges between 20% and 70%. A study conducted by Keropian and Murphy, comprising 273 patients presenting with sleep-disordered breathing (SDB) visiting the author's clinic, revealed snoring to be the most common presenting complaint among the SDB subjects, marked in all 273 cases. The results of 273 PSGs revealed an elevated (96%) prevalence of OSA in the study.[8] In a study conducted by Vulli *et al*., OSA was seen to be prevalent in 85.71% of the suspected cases of OSA, with 28 out of 30 OSA cases belonging to moderate-to-severe category.[9] Thus, enough evidence is obtained to explain the high OSA prevalence observed in the current study.

In the present study, majority of cases, 83.08% (*n* = 54), had a high OSA risk according to the SBQ. On the other hand, a smaller group of 11 cases, representing 16.92% of the total, was determined to have a low likelihood of OSA according to the questionnaire. ROC curve analysis of STOP-BANG score in the current study showed that cutoff score of >5 can be used to predict high risk of OSA [Figure 1]. At this cutoff, the sensitivity was 91.23% with a confidence interval of 80.7%–97.1%, and the specificity was 75% with a confidence interval of 34.9%–96.8%, with *P* = 0.0185 indicating statistical significance. The PPV was 96.3% (87.3%–99.5%), and the NPV was 54.5% (23.4%–83.3%). The overall diagnostic accuracy was 89.23%, with *P* = 0.016 from the Chi-square test. The result was in concordance with previous studies. Vulli *et al*. conducted a study which aimed to compare the predictive abilities of ESS and STOP-BANG scores for OSA. The STOP-BANG score showed the highest sensitivity of 96.66% for predicting OSA, with a specificity, PPV, and NPV of 40%, 90.62%, and 66.66%, respectively.[9] In a similar study conducted by El-Sayed, results revealed STOP-BANG to have high sensitivity (97.55%) with a low

specificity (26.32%) in predicting OSA. PPV and NPV as revealed by the study were 93.43% and 50%, respectively.[10] Pataka *et al*. conducted a comparative study among the screening scores of OSA, including the STOP-BANG score and ESS. The results found STOP-BANG to have a high sensitivity (97.60%) with a low specificity (12.70%) with a PPV of 83.30% and an NPV of 45% in the suspected OSA patients.[11]

When ESS was used to evaluate the risk of OSA, results revealed that 64.62% (n = 42) of cases had a high risk of OSA. The remaining 23 cases, which accounted for 35.38%, were assessed as having a low risk of OSA based on their ESS scores. For the ESS score, ROC curve analysis in the current study showed that cutoff score of ≥11 can be used to predict high risk of OSA [Figure 2]. At this cutoff, the sensitivity was 70.18% with a confidence interval of 56.6%–81.60%, and the specificity was 75% with a confidence interval of 34.9%–96.8% with P = 0.0014, also indicating statistical significance. The PPV was at 95.2% (83.8%–99.4%), and the NPV was significantly lower at 26.1% (10.2%–48.4%). The overall diagnostic accuracy was 70.77% from the McNemar test. In the study, the mean ESS scores varied among different OSA severity levels but did not show a statistically significant association, as evidenced by P = 0.404. This is explainable by the fact that ESS is a predictor scale for subjective sleepiness, which may occur due to various other causes and not only OSA. The result was comparable to a similar study conducted by El-Sayed, where ESS was found to have a specificity of 75% but with a sensitivity of 72.55% in screening OSA in the suspected group of patients.[10] Another study conducted by Vulli *et al*. that aimed to compare the two scores for OSA screening found out ESS to have a sensitivity of 73.33% with a specificity of 60% among the study participants in OSA screening. The score had a PPV and NPV of 91.66% and 27.27%, respectively.[9]

The current study established that the SBQ outperformed the ESS questionnaire in identifying OSA cases. Taking into account the at-risk patients for OSA using the two questionnaires, this study revealed that the prevalence of OSA in the study group was 87.69%; 83.08% were classified as having a high OSA risk by the SBQ (score >5), and 64.62% were classified as being at high risk of the same by the ESS (score ≥11). When compared to the SBQ (91.23%), the ESS had the low sensitivity (70.18%) with similar specificity (75%) for predicting OSA, thus making STOP-BANG a better tool over ESS for OSA screening. This is explainable by the fact that the ESS is a commonly used questionnaire to assess subjective excessive daytime sleepiness, which is a diagnostic criterion for OSA but can also result from a variety of secondary conditions. The results were similar to those of previous studies. In a study conducted by Zheng *et al*. including 1671 patients with suspicion of OSA, the AUC evaluated by STOP-Bang were 0.724, 0.703, and 0.712, and those of ESS were 0.632, 0.634, and 0.695, with AHI cutoffs of ≥5, ≥15, and ≥30 events/h, respectively. The results revealed the STOP-BANG score to be a better predictor of OSA, being highly sensitivity but less specific than ESS.[12] The study conducted by El-Sayed revealed STOP-BANG to have high sensitivity (97.55%) with a low specificity (26.32%) in predicting OSA. ESS was found to have a specificity of 75% but with a low sensitivity (72.55%) in screening OSA. Hence, these studies conclude that SBQ is a better tool than ESS to screen patients for OSA.[10] In a similar study, Pataka *et al*. found out STOP-BANG to have a high sensitivity (97.6%) in OSA screening over ESS (50%). With a specificity of 12.70% and 67%, respectively, the study concluded STOP-BANG to be a better screening questionnaire of OSA over ESS.[11]

The present study is among the few Indian studies to compare the screening questionnaires of OSA to provide a better tool for its detection, in this geographical location. We compared the STOP-BANG scale with ESS and found that both tests showed relatively high diagnostic accuracy and predictive values, particularly the STOP-BANG score, which outperformed the ESS score in most metrics. The study had a few limitations also. The participants included in the study were enrolled from the OPD based on high suspicion of OSA. This might lead to selection bias, thus affecting the OSA prevalence in the general population. A small sample size of 65 patients may have affected the generalizability of the results. To further validate the findings, multicenter studies, including a bigger sample size, are necessary.

Conclusion

The present study was able to provide valuable insights into OSA screening. Conducted with the aim of comparison of the two commonly used OSA screening questionnaires, we found out that the STOP-BANG score and ESS, both, had comparatively good predictive and diagnostic accuracy, with the STOP-BANG score surpassing the ESS score in the majority of measures. Considering the high global burden of undiagnosed OSA, there is a need to upregulate the screening for OSA followed by appropriate treatment measures. This would improve sleep quality, eliminate apneas and hypopneas, and normalize SpO_2 levels to reduce the risk of future adverse health outcomes.

Ethics statement

The study protocol was reviewed and approved by the Institutional Ethics Committee, Government Medical College and Hospital (GMCH), Sector 32, Chandigarh, India, with Document Approval Number GMCH/IEC/774R/2022/198.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, *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;13:479-504.
2. Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: A population health perspective. Am J Respir Crit Care Med 2002;165:1217-39.
3. Johns MW. A new method for measuring daytime sleepiness: The Epworth sleepiness scale. Sleep 1991;14:540-5.
4. Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, *et al*. STOP questionnaire: A tool to screen patients for obstructive sleep apnea. Anesthesiology 2008;108:812-21.
5. Chung F, Abdullah HR, Liao P. STOP-bang questionnaire: A practical approach to screen for obstructive sleep apnea. Chest 2016;149:631-8.
6. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MS, Morrell MJ, *et al*. Estimation of the global prevalence and burden of obstructive sleep apnoea: A literature-based analysis. Lancet Respir Med 2019;7:687-98.
7. Quan SF, Gillin JC, Littner MR, Shepard JW. Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research editorials. Sleep 1999;22:662-89.
8. Keropian B, Murphy N. The prevalence of OSA in snorers presenting with various chief complaints: A pilot study. Cranio 2014;32:217-8.
9. Vulli V, Balla P, Vasu P. Predictive ability of STOPBANG scale and Epworth Sleepiness Scale in identifying obstructive sleep apnoea. IOSR J Dent Med Sci 2019;18:1-7.
10. El-Sayed IH. Comparison of four sleep questionnaires for screening obstructive sleep apnea. Egypt J Chest Dis Tuberc 2012;61:433-41.
11. 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;15:776-81.
12. Zheng Z, Zhang Y, Chen M, Chen X, Li C, Wang C, *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.