Otolaryngology–Head and Neck Surgery (2005) 133, 966-971

ORIGINAL RESEARCH

The Association of Tongue Scalloping With Obstructive Sleep Apnea and Related Sleep Pathology

Todd M. Weiss, MD, Strahil Atanasov, MD, and
Karen H. Calhoun, MD, FACS, Springfield, Illinois; Galveston, Texas; and Columbia, Missouri

OBJECTIVE: The association between OSAS and patient history and physical exam findings is previously established; however, to our knowledge there are no studies that evaluate the role of tongue scalloping as a reliable clinical indicator for OSA, snoring, or the presence of other sleep pathology as evidenced by polysomnography. This study evaluates the hypothesis that such an association exists.

SUBJECTS AND METHODS: Sixty-one otolaryngology clinic patients were evaluated by history and physical exam for the presence and degree of tongue scalloping, snoring, and other previously established clinical indicators for sleep-disordered breathing and obstructive apnea. Twenty-five of the 61 study patients were additionally evaluated by overnight polysomnography to provide conclusive diagnosis of sleep pathology. The degree of tongue scalloping was graded from 0 to 3 and its significance as a screening, diagnostic, and predictive factor for sleep pathology was then statistically determined.

RESULTS: Twenty-seven patients (44%) had known or newly documented OSA and 47 (77%) had a history of snoring. Twenty-seven patients (44%) had some degree of tongue scalloping (1-3) and 74% of these patients were male. The presence of any degree of tongue scalloping (grade 1-3) in patients with known or newly documented OSA showed sensitivity, specificity, PPV, and NPV of 52%, 68%, 70%, and 50% respectively. The presence of tongue scalloping in patients with either known snoring history or newly documented snoring showed sensitivity, specificity, PPV, and NPV of 47%, 64%, 81%, and 26% respectively. Presence of tongue scalloping was 71% specific for abnormal sleep efficiency (<85%), 70% specific for abnormal AHI (>5), and 86% specific for nocturnal desaturation >4% below baseline. Presence of tongue scalloping also showed PPV of 67% for abnormal AHI, 89% for apnea or hypopnea, and 89% for nocturnal desaturation. Presence and severity of tongue scalloping showed positive correlation with increasing Mallampati and modified Mallampati airway classification.

CONCLUSIONS: In high-risk patients we found tongue scalloping to be predictive of sleep pathology. Tongue scalloping was also associated with pathologic polysomnography data and abnormal Mallampati grades. We feel the finding of tongue scalloping is a useful clinical indicator of sleep pathology and that its presence should prompt the physician to inquire about snoring history.

EBM RATING: C

Sleep pathology among Americans occurs frequently and has high associated morbidity. Among the most prevalent sleep disorders are habitual snoring and obstructive sleep apnea syndrome (OSAS), where both result in significant social and physical consequences. The prevalence of OSAS among US adults varies depending upon the diagnostic criteria used. A large population-based study by Young et al[1] recently reported the prevalence of OSAS (RDI > 5, self-reported daytime somnolence) in the general population to be 4% in men and 2% in women, similar to that of diabetes mellitus or asthma. Sleep-disordered breathing is even more common in the elderly population where one community-based sample demonstrated prevalence between 44% (RDI > 20) and 81% (RDI > 5).[2] Habitual

From the Department of Otolaryngology, Southern Illinois University (Dr Weiss), Department of Neurology, University of Texas Medical Branch (Dr Atanasov), and Department of Otolaryngology–Head & Neck Surgery, University of Missouri (Dr Calhoun).

Reprint requests: Karen H. Calhoun, MD, FACS, Department of Otolaryngology–Head & Neck Surgery, University of Missouri Medical School, Columbia, MO 65212.

E-mail address: calhounk@missouri.edu.

0194-5998/$30.00
doi:10.1016/j.otohns.2005.07.018

snoring is an upper airway resistance problem common among middle-aged Americans where as many as 44% of men and 28% of women are known to be affected.[3] This too was reported to have increasing prevalence with age.

Patients presenting with sleep disorders are clinically evaluated using history and physical exam, while definitive diagnosis can only be made using polysomnography or "sleep study." Unfortunately for these patients, the common symptoms are not reliable clinical indicators for OSAS. Clinical impression alone has poor (50%-60%) sensitivity and low (63%-70%) specificity for detecting OSAS, and patients with objectively diagnosed sleep-disordered breathing are believed to have problematic symptoms of insomnia only 50% of the time.[4] Polysomnography is expensive, inconvenient and uncomfortable for the patient, and time consuming for both the patient and hospital staff. This method of diagnosis can therefore proceed only after thorough clinical justification so that the patient will give consent and the hospital will be reimbursed.

Subjective adjuncts to diagnosis aid clinicians in this respect and have been employed using patient-directed questionnaires that assign a numerical score to quantify the severity of the symptoms. An example frequently used among sleep labs and clinics is the Epworth Sleepiness Scale (ESS). This 8-part questionnaire consistently assigns high scores to patients with apnea and has even demonstrated correlation between disease severity and score elevation.[5] In efforts to further validate the polysomnographic workup of patients with subjective symptoms of excessive daytime sleepiness and sleep deprivation, objective data has been sought.

Physical markers are reliable and easy to obtain by exam, and because the findings are objective, they maintain a universal standard for comparison. Markers previously described to be associated with or suspicious for sleep pathology and/or difficult oral-tracheal intubation include tonsillar/adenoid hypertrophy, macroglossia with glossopalatal disproportion, large neck circumference, obesity measured by body mass index (BMI), inferior displacement of the hyoid bone documented by increased hyoid-mental distance, increased thyroid-mental distance, and the Mallampati airway classification scale.[6] Although objective indices, when abnormal, have been loosely associated with difficult oral-tracheal intubation, OSAS, and snoring disorder, no individual finding or score is a reliable indicator of diagnosable sleep pathology, just as no individual symptom or sign can predict with accuracy the presence or absence of OSA.[7] Attempts are therefore being made to develop algorithms and clinical prediction models that place significant objective findings in combination so as to produce reliable and sensitive screening results.[8] A highly sensitive method is necessary if physicians are to confidently and accurately recommend patients for polysomnography.

Clinic patients suffering from either diagnosed OSAS or other sleep-related disorders that posses high-risk physical markers for sleep apnea have been noted to repeatedly exhibit some degree of tongue scalloping. We define tongue scalloping as the multiple lateral glossal indentations resulting from molar compression. This condition is secondary to either glossopalatal disproportion alone or in combination with macroglossia. To our knowledge there are no previous studies describing this clinical finding or its relationship to the various types of sleep pathology. The purpose of this study is to broaden the spectrum of clinical parameters used to justify polysomnography by defining a physical marker and examining its relationship to different forms of sleep pathology. Our study investigates the association between tongue scalloping and OSAS, habitual snoring, and other pathologic sleep parameters as evidenced by polysomnography with the hypothesis that such an association exists.

METHODS

Subjects

Sixty-seven otolaryngology clinic patients aged between 20 and 82 years (mean age 50.6 $\pm$ 14.9) were asked to participate in the study. Patient selection was indiscriminate and patients were randomly selected from one of four otolaryngology specialty clinics including head and neck, otology, sleep, and private faculty clinic. Patients over the age of 18 and not associated with the Texas Department of Corrections prison hospital were accepted for the study without regard to their history of OSA. Thirty-one of the 67 patients were scheduled for polysomnography prior to the time that they were enlisted in the study. These patients were evaluated in the sleep lab according to the study inclusion criteria. An all-night comprehensive sleep study was performed and all relevant polysomnographic data was obtained. All subjects consented verbally and the Department of Otolaryngology along with the Internal Review Board at the University of Texas in Galveston approved the study.

Measures and Procedures

All patients were evaluated by oropharyngeal physical exam for the presence and degree of tongue scalloping (Fig 1). The degree of tongue scalloping was given a numerical score between 0 and 3 based upon the examiner's findings and was classified as either absent, mild, moderate, or severe, respectively. The standard criteria used to select the severity of scalloping were complete absence, scalloping evident but not pronounced, scalloping pronounced but resolved with tongue protrusion, and scalloping pronounced and unresolved with tongue protrusion. Patient histories with respect to snoring and previous diagnosis of OSA were recorded. History of snoring was considered positive if verified by either the subject or the subject's spouse, and the presence of snoring was documented by history and/or polysomnography. All patients were individually evaluated for the following established or suspected clinical indicators of sleep pathology: obesity (BMI $\geq$ 30), neck circumference, thyroid-mental distance measured from the thyroid notch to

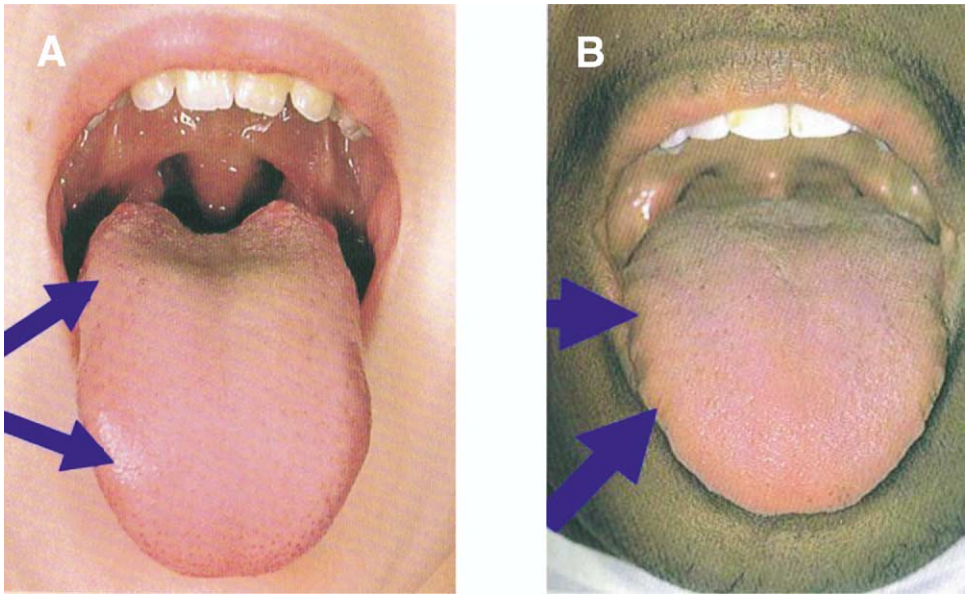

Figure 1 **(A)** Lateral glossal margin in a normal patient **(B)** in patient with grade III tongue scalloping evident during tongue protrusion.

the mental prominence, hyoid-mental distance measured from the hyoid bone to the mental prominence, and tonsil grade. Tonsil grade provided simplistic assessment of oropharyngeal occlusion and was graded (0-4) according to the parameters set by Friedman et al.[9] All physical measurements were recorded and performed with the patients seated upright facing the examiner holding their heads in natural position. Indices for difficulty of intubation were assessed by examination of the oropharynx and included the traditional Mallampati score plus a modified Mallampati (MMP) score described by Friedman et al.[9] Traditional Mallampati signs designate airway class I-IV, where classes I and II indicate no difficulty, and classes III and IV indicate moderate and severe difficulty, respectively.[10] The MMP retains the same airway class designations and associated difficulties as the traditional Mallampati, the difference being an absence of tongue protrusion during exam, maintaining a more natural, resting lingual approximation. All patients completed the standard ESS and a 20-part questionnaire designed by the authors concerning signs and symptoms of daytime sleepiness. This questionnaire was scored from 0 to 20 where a score of 20 indicated maximal suspicion for sleep pathology.

Polysomnography

Thirty-one patients were additionally evaluated by a fully attended overnight hospital-based polysomnography to provide conclusive diagnosis of sleep pathology. Complete sleep data sets from only 25 of these patients was obtained and used for the purpose of this study. The remaining 6 patients had either incomplete sleep data or records that were inaccessible during the study review. The nocturnal polysomnography included the recording of EEG (C3/A2 and O2/A1), EOG, EKG, EMG (chin and tibialis anterior muscle), breath sounds (tracheal microphone), SaO_2 (oxymeter), chest-wall and abdominal movements (respitrace), and nasal/oral airflow (thermistor). A 6-hour minimum duration was required for diagnosis. Data obtained included sleep efficiency, number and duration of apneas/hypopneas, apnea-hypopnea index (AHI), sleep onset latency, periodic leg movement index (PLMI), minimum and average oxygen saturation, number of desaturations greater than 4% below baseline, and final diagnosis. Apnea was defined as cessation of airflow greater than 10 seconds. Hypopnea was defined as a recognizable, transient reduction, but not complete cessation of breathing greater than 10 seconds and associated with desaturation greater than 4% below baseline, a cortical arousal, or both. The AHI represents the sum of all apneas and hypopneas per hour. OSAS was defined as AHI greater than 5 as all patients undergoing polysomnography were symptomatic.[11]

Statistical Analysis

All data was recorded and stored in a Microsoft Excel spreadsheet for statistical analysis. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated using a 2-way contingency table. Statistical correlation was assessed by *P* value using 95% confidence intervals. The likelihood of possessing tongue scalloping in conjunction with specific disease states was interpreted by relative risk. Bivariate distribution figures were generated using an X-Y scatter plot and analyzed by a line of best-fit regression to calculate the Pearson coefficient (r) and the coefficient of determination (r^2). Individual regression line equations were recorded showing correlation between the severity of tongue scalloping and increasing degree of pathologic findings. Regression lines were of either binomial or trinomial order and therefore a single correlation coefficient was not known. The relationship between tongue scalloping with OSA, snoring, obesity,

Table 1
Statistical analysis of patient physical exam findings and questionnaire results

	Patient age	BMI (kg/m^2)	Neck circumference (cm)	ESS	Degree of scalloping	Tonsil grade	Mallampati	MMP	Total questionnaire "yes"
Mean	50.6	33.8	43.5	9.5	0.61	1.13	2.2	2.3	7.9
Std dev	14.9	9.5	8.1	6.3	0.82	1.22	1.0	1.2	4.9
Range	20-84	18.3-68.9	32.0-67.3	0-24	0-3	0-4	1-4	1-4	0-17

sleep efficiency, AHI, and other polysomnographic variables was then determined.

RESULTS

The study group for which all data was obtained consisted of 61 patients (36 males and 25 females) with an average study age of 50.62 (± 14.87) years. The age range fell between 20 and 84 years. Table 1 shows statistical analysis of physical exam findings and questionnaire results.

The overall prevalence of any degree tongue scalloping (grade 1-3) among all study patients was 44.26% (n = 27). Fifty-five percent of males and 28% of females showed evidence of tongue scalloping, while 74% of all patients with tongue scalloping were male. Tongue scalloping provided evidence of OSA, habitual snoring, and obesity to varying degrees. Results are displayed in Table 2. Additionally, the results of overnight polysomnography were obtained on 25 patients. The association of tongue scalloping and abnormal polysomnographic data is shown in Table 3.

Seventy-five percent of all study patients (n = 46) had documentation of OSA disease status. Statistical relationships involving study variables specific to OSA were obtained from this patient group only. Variables that did not require known OSA status were analyzed using the entire patient pool. A control group of 29 patients within the total study population did not have documented OSA. In 7 patients this was demonstrated by negative polysomnography findings.

Forty-four percent (n = 27) of all study patients had either previous or newly diagnosed OSA. Tongue scalloping was present in 70% of all patients with known or newly documented OSA. Tongue scalloping was present in 58% (n = 17) of patients without documented OSA (grades 1 and 2 only). However, of the 7 patients in the control group who underwent polysomnography, only one was found to have any evidence of tongue scalloping, which was grade 1. The presence of any degree of tongue scalloping (grade 1-3) in patients with known or newly documented OSA showed sensitivity, specificity, PPV, and NPV of 52%, 68%, 70%, and 50% respectively (P = 0.28) (RR = 1.4).

Seventy-nine percent (n = 48) of all study patients had a documented presence of snoring. Tongue scalloping was present in 81% of documented snorers. The presence of any degree of tongue scalloping in patients with documented snoring showed sensitivity, specificity, PPV, and NPV of 46%, 62%, 81%, and 24% respectively. The presence of any degree of scalloping was 89% predictive for both the presence of nocturnal apnea and oxygen desaturation greater than 4% below baseline. Tongue scalloping was 67% predictive for AHI greater than 5, which met our studies criteria for OSA.

All patients with grade 3 tongue scalloping (n = 3) were diagnosed with OSA. Patients with any degree of tongue scalloping were 1.3 times more likely to have an AHI greater than 5, 1.4 times more likely to have a diagnosis of OSA or nocturnal desaturations, 1.5 times more likely to have sleep efficiency less than 85%, and 2 to 3 times more

Table 2
Tongue scalloping as clinical evidence of OSA, habitual snoring, and obesity

	Scalloping and OSA n = 46	Scalloping and snoring n = 48	Scalloping and obesity n = 61
Sensitivity	51.85 [38.0-63.0]	45.83 [39.8-53.1]	45.71 [33.8-57.2]
Specificity	68.42 [48.8-84.8]	61.54 [16.0-65.5]	57.69 [41.7-73.1]
PPV	70.00 [51.3-85.5]	81.48 [63.6-85.0]	59.26 [43.8-74.1]
NPV	50.00 [35.6-61.9]	23.52 [06.7-27.5]	44.12 [31.9-55.9]
P value	0.280	0.480	0.997
RR	1.400	0.874	1.060

Values bracketed indicate 95% confidence interval.

Table 3
Tongue scalloping and relation to sleep pathology detected by polysomnography (n = 25)

	Sleep efficiency (<85%)	AHI > 5	Presence of apnea	Nocturnal desaturation (>4% below baseline)
Sensitivity	45.45 [21.3-67.0]	42.86 [23.3-57.9]	40.00 [28.3-44.7]	44.44 [29.8-49.7]
Specificity	71.43 [52.5-88.3]	72.73 [47.8-91.8]	80.00 [33.4-98.9]	85.71 [48.0-99.2]
PPV	55.56 [26.1-81.8]	66.67 [36.2-90.0]	88.89 [63.0-99.4]	88.89 [59.5-99.4]
NPV	62.50 [45.9-77.3]	50.00 [32.9-63.1]	25.00 [10.4-30.9]	37.50 [21.0-43.4]
P value	0.650	0.699	0.755	0.344
RR	1.481	1.330	1.185	1.422

Values bracketed indicate 95% confidence interval.

likely to have a high-risk airway class measured by Mallampati and MMP scales respectively.

Results of the bivariate distribution figures using X-Y scatter plots demonstrated patients between the ages of 40 and 50 years were found to have the most severe degrees of tongue scalloping, with a peak at 45 years of age. There were, however, no statistically significant trend lines from this data.

Higher degrees of tongue scalloping correlated with higher airway class designation using Mallampati (r^2 = 0.19) and MMP (r^2 = 0.21). The association between tongue scalloping and a Mallampati class 3-4 airway has 63% sensitivity, 68% specificity, 56% PPV, 74% NPV, and RR 2.1 (P = 0.04). The association between tongue scalloping and the modified Mallampati class 3-4 airway has 72% sensitivity, 75% specificity, 67% PPV, 80% NPV, and RR 3.2 (P = 0.001). The degree of tongue scalloping showed no linear correlation to increased number of nocturnal desaturations, increased Epworth Sleepiness Scale score, increased number of symptoms and risk factors as assessed by questionnaire, increased neck circumference, or increased tonsil grade.

DISCUSSION

The purpose of this study was to determine the relationship, if any, between the clinical exam finding of tongue scalloping and objective evidence of sleep pathology. Obstructive sleep apnea is defined by an abnormal AHI, yet the exact value of AHI used to define OSA presents as a well-recognized discrepancy within the current literature. Recognizing this inconsistency, our study pursued a broader spectrum of pathologic indices and considered multiple polysomnographic parameters for evidence of sleep-disordered breathing, including the mere presence of apnea and/or hypopnea, nocturnal desaturation, and sleep efficiency. We considered an AHI greater than 5 necessary for the presence of OSAS because all patients undergoing polysomnography were symptomatic. We also considered any of the following findings abnormal: the presence of any apnea, sleep efficiency less than 85%, and nocturnal desaturation greater than 4% below baseline. Our results indicate that tongue scalloping was predictive for the presence of many of these individual parameters within our patient population, all of which indicate sleep pathology to varying degrees. The presence of tongue scalloping was 70% predictive for the total number of patients with either previous or newly diagnosed OSA. This is similar to the 67% predictive value tongue scalloping had for the patients in our study who underwent polysomnography and had an AHI greater than 5, consistent with our criteria for OSAS. We believe the reproducibility of the predictive value in two different patient groups enforces its validity.

An overwhelming majority of patients in the study had grade 1-2 tonsils or history of tonsillectomy, and therefore higher degrees of tongue scalloping did not correlate with increased tonsil grade. This is largely due to our adult patient population, where acute or chronic tonsillar hypertrophy was not clinically relevant. Enlarged neck circumference is a well-known predictor of difficult intubation and has been shown a risk factor for OSA. Our study revealed no clear association between enlarged neck circumference and tongue scalloping, and no significant trend toward increased neck circumference and high-grade scalloping. Neck circumference, although a logical risk for upper airway obstruction, likely plays no role in the pathophysiology of tongue scalloping. Most patients with severe scalloping (grade 3) were between 40 and 50 years of age. We currently offer no physiologic explanation for this; however, statistically this age range included both the mean and mode for the study.

The 20-part questionnaire was found useful in two respects: determining signs and symptoms of daytime somnolence and ruling out co-morbid conditions such as hypothyroidism; however, it showed no significant linear correlation to the presence of documented sleep pathology. High ESS scores also demonstrated no significant correlation to increasing grades of tongue scalloping. Patients with the highest ESS scores (18-24) demonstrated no evidence of tongue scalloping, distorting the reliability of the trend line. However, this result can be expected without predictive values equal to or higher than 90%.

Study data showed the presence of tongue scalloping to actually decrease the overall patient risk for snoring, although scalloping was associated with a high PPV for detecting snorers. This apparent inconsistency is due to the large majority of study patients with documented snoring history. The ratio of patients with documented snoring history to patients with tongue scalloping was disproportionately large so that the presence of snoring falsely indicates a decreased likelihood of having tongue scalloping, while the presence of tongue scalloping indicates high predictability for patients who snore. This suggests why the relative risk is less than 1. A more accurate and less misleading result would have been obtained if more patients with no history of snoring were included in the study. Therefore, results from the subset of patients without snoring history should be considered more reliable in this instance, where the absence of documented snoring history was 62% specific for patients to have no degree of tongue scalloping, and the absence of tongue scalloping was only 24% predictive for nonsnorers.

An explanation for the relatively high predictive values in this study is the bias in patient population toward preexisting sleep pathology. Despite attempts to diversify the patient pool, a large number of patients were recruited from otolaryngology sleep clinics. Also, all patients undergoing polysomnography were scheduled secondary to legitimate sleep-altered symptoms and diagnostic workup protocols. No patient was scheduled on the basis of this study alone. Only 31 of 67 patients underwent overnight PSG. Of the remaining 36 patients, OSA status was known in only 10. Increasing the number of patients with polysomnography-diagnosed OSA disease status would increase the study power and thus the overall accuracy of the results.

Another contributing bias not necessarily related to our specialty was the large number of patients who snore that were included in the study. The prevalence of snoring in our study was 79%. Although this number may not be consistent to reports within the general population, it may perhaps be more accurate as all snoring documentation was largely provided by the patient's spouse and/or polysomnography, both considered to be a more accurate source than the patient alone.

High prevalence of sleep-related breathing disorders within our patient population is the reason for high positive predictive values reported by this study. The PPV increases as the prevalence of the disease increases. Therefore it can be inferred that the finding of tongue scalloping in patients considered high risk for OSA is largely suggestive of the disorder. It is known that the prevalence of sleep disorders within the general population is relatively high. In this study alone the prevalence of OSA and habitual snoring were 47% and 79% respectively. Our study found tongue scalloping to be 70% predictive for diagnosis of OSA. We believe that the presence of tongue scalloping in patients at increased risk for sleep pathology as assessed by symptoms and clinical exam is predictive for the polysomnographic diagnosis of varying sleep pathologies, including but not limited to apnea and OSAS. We therefore feel the finding of tongue scalloping in any patient is a useful clinical indicator of sleep pathology and that its presence should prompt the physician to inquire about snoring history. Until the actual prevalence of sleep pathology within the general population is known, tongue scalloping should alert the examining clinician to have a high index of suspicion for sleep disturbance.

REFERENCES

1. Young T, Palta M, Dempsey J, et al. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 1993;328: 1230–5.
2. Ancoli-Israel S, Kripke DF, Klauber MR, et al. Sleep-disordered breathing in community-dwelling elderly. Sleep 1991;14:486–95.
3. Bailey BJ. Head and neck surgery–otolaryngology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 579.
4. Skomro RP, Kryger MH. Clinical presentations of obstructive sleep apnea syndrome. Prog Cardiovasc Dis 1999;41(5):331–40.
5. Johns MW. A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. Sleep 1991;14:540–5.
6. Hiremath AS, Hillman DR, James AL, et al. Relationship between difficult tracheal intubation and obstructive sleep apnea. Br J Anaesth 1998;80:606–11.
7. Herer B, Roche N, Carton M, et al. Value of clinical, functional, and oximetric data for the prediction of obstructive sleep apnea in obese patients. Chest 1999 (Dec);116(6):1537–44.
8. Kirby SD, Eng P, Danter W, et al. Neural network prediction of obstructive sleep apnea from clinical criteria. Chest 1999 (Aug); 116(2):409–15.
9. Friedman M, Tanyeri H, La Rosa M, et al. Clinical predictors of obstructive sleep apnea. Laryngoscope 1999 (Dec);109(12):1901–7.
10. Mallampati S, Gugino L, Desai S, et al. A clinical sign to predict difficult tracheal intubation: a prospective study. Can Anesth Soc J 1985;32:429–34.
11. Loube DI, Gay PC, Strohl KP, et al. Indications for positive airway pressure treatment of adult obstructive sleep apnea patients. Chest 1999 (March);115(3):863–6.