CHEST Online Supplement

Digital Morphometrics

A New Upper Airway Phenotyping Paradigm in OSA

Richard J. Schwab, MD; Sarah E. Leinwand, MPH; Cary B. Bearn, MS; Greg Maislin, MS; Ramya Bhat Rao, DMD; Adithya Nagaraja, BA; Stephen Wang, BS; and Brendan T. Keenan, MS

CHEST 2017; 152(2):330-342

 DOI: j.chest.2017.05.005

e-Appendix 1.

SUPPLEMENTAL METHODS

Study Population

Subjects were recruited from the University of Pennsylvania Sleep Center. There were no exclusions based on BMI or previous medical history; subjects were excluded if they refused an overnight sleep study. Apneics were defined as having an AHI≥10 events/hour and controls as an AHI<10. Upon recruitment, gender, race, age, height, and weight were recorded, BMI was calculated and morphometric photographs were obtained. A description of the photographs and measurements obtained is presented in **Tables 1 and 2** in the manuscript.

Overnight Polysomnography

Subjects underwent either an overnight polysomnograph at Penn Center for Sleep Disorders [n=787 (91.5%)] or a home study with an Embletta Gold portable monitor (Natus Medical Incorporated) [n=73 (8.5%)], depending on the research protocol and/or their insurance coverage. Each study was scored by a trained sleep technologist and reviewed by certified sleep physicians according to the American Academy of Sleep Medicine scoring methods [1].

Laser Ruler

Digital photographs were obtained using an intraoral laser ruler and digital camera (**Figure 1**). The laser ruler was composed of a right angle beamsplitter and mirror aligned such that two parallel beams project forward a known distance apart. A camera was attached to the laser ruler (**Figure 1**) to allow for digital photographs to capture the projected laser beams adjacent to measures of interest. The known distance between the lasers, either 1.0 or 1.5 centimeters depending on the specific device being used, can then be used to calculate measures directly from the photograph. The bright centroid of each laser was selected using the wand plugin on ImageJ (1.4.3.67 http://rsb.info.nih.gov/ij/ NIH, USA). The (x, y) coordinates for the centroids of each laser were noted and the distance between the centroids [(x1, y1) and (x2, y2)] was calculated using the Pythagorean distance formula [$distance = \sqrt{(x2 - x1)^2 + (y2 - y1)^2}$]. This distance, or conversion factor, was then used to convert pixel measures taken from the photographs into centimeters for all measures using the formula:

$$Measurement\ in\ centimeters = Measurement\ in\ Pixels \times \frac{Distance\ between\ lasers\ in\ cm}{Distance\ between\ lasers\ in\ pixels}$$

Morphometric Photographs

For each morphometric photograph, subjects were seated with their head in a neutral position, such that their line of sight was parallel to the floor [2]. The camera and laser ruler were placed between 35 and 50 centimeters from the subject. Subjects were instructed to open their mouths maximally for a series of 4 intraoral photographs (**Table 1**) in the frontal or profile (side of the face) position: tongue in the mouth without phonation (frontal photo), tongue extended (frontal photo), tongue extended (profile photo), tongue depressed without phonation (frontal photo) (**Figure 2**). The length, width, and area measurements taken from the photographs are described in **Table 2** and shown in **Figure 2**.

We also calculated categorical measures of pharyngeal airway visibility (visible/not visible), modified and standard Mallampati, evidence of tongue ridging, tonsil hypertrophy grade, and pharyngeal narrowing grade for each subject (see **Figure 3**). Modified Mallampati was scored using the tongue in the mouth photograph without phonation (**Figure 3**), whenever the palate was clearly visible regardless of the inclusion of lasers. Pharyngeal airway visibility was derived from this measure, classified as visible if the subject had a modified Mallampati Class of I or II and not visible for Classes III and IV (**Figure 3**). Tongue ridging severity was scored if the tongue edges could be clearly visualized in the frontal intraoral photographs. Tongue ridging was classified as evident if it was observed in any photographs [3] (**Figure 3**). The tonsil hypertrophy grade was scored if full visibility of the airway was achieved in any of the frontal intraoral photographs and did not require the inclusion of the lasers. Tonsil hypertrophy was graded on a standard scale from 0 to 4: 0) when the tonsils were not visible because either they were hidden within the tonsillar fossa or otherwise removed; 1) when the tonsils were seen within the pillars; 2) when the tonsils extended to the pillars; 3) when the tonsils extended laterally beyond the pillars; and 4) when the tonsils extended to the midline [4, 5]. Given the low prevalence of tonsil grade 4, we combined grades 3 and 4 in the analyses. The pharyngeal narrowing grade was also scored if full visibility of the airway was achieved in any of the frontal intraoral photographs and did not require the inclusion of the lasers. Pharyngeal narrowing was graded on a scale from 1 to 4: 1) the palatopharyngeal arch intersects at the edge of the tongue; 2) the palatopharyngeal arch intersects at 25% or more of the tongue width; 3) the palatopharyngeal arch intersects at 50% or more of the tongue width; and 4) the palatopharyngeal arch intersects at 75% or more of the tongue width (**Figure 3**).

Validation and Reproducibility

To assess accuracy and reliability of the laser ruler, we used the laser ruler to calculate the diameter and area of a circle with a 2 inch diameter in 10 photos taken at a distance of 40 cm. To examine the effect of the distance between the camera and the object on the reproducibility of measurements, we used the laser ruler to measure the dimensions of the same circle in 1 cm increments at distances from 35 to 50 cm (representing the range of distances from which photos on subjects were taken). A laser range finder (Bosch DLR165) with an accuracy of 0.15 cm was used to measure the distance between the circle and the body of the digital camera. Measurements of the circle taken with the digital ruler were evaluated based on the coefficient of variation for both the calculated diameter and area. Low coefficients of variation indicate reproducible measures. We also calculated the average difference between the estimated and actual circle diameter and area measures.

To test the reproducibility and validity of the photographic procedure and subject positioning we took two photographs of 10 subjects (convenience sample), within a 1 month interval. Our analysis method was further validated by repeating the analysis on the same photograph for 10 random subjects. We examined the reproducibility and validity of the photographs and analysis methods (see more detail below), with results presented in **e-Tables 1-3,** below. All pictures were taken and analyzed by 2 trained technologists (SEL and CB) using the same validated protocol. Both technologists were blinded with respect to apnea status.

Statistical Methods

Absolute and distance error were evaluated using percent coefficients of variation (CV) $[\%CV = (\frac{Standard\ Deviation}{Mean} \times 100)]$. Intra-class correlation coefficients (ICCs) were used to evaluate reproducibility of the continuous anatomical measures and percent agreement was used to show reproducibility of the more subjective categorical anatomical measures. An ICC >0.70 or percent agreement >70% were considered acceptable levels of reproducibility. Continuous sample characteristics were summarized using means and standard deviations and were compared between groups using t-tests. Categorical covariates were summarized using frequencies and percentages and compared between groups using chi-square tests or Fisher's exact tests.

We examined the relationship between our photographic variables and OSA status using logistic regression models. Results are presented using odds ratios (ORs) and 95% confidence intervals (CIs) associated with a 1 standard deviation (SD) increase in continuous variables or compared to the indicated reference group for categorical variables. Similarly, we examined the association between photography variables and OSA severity using linear regression models with ln(AHI + 1) as the outcome, log transformed to allow for parametric analyses. Results of these models are presented as mean changes in ln(AHI + 1) (natural log) and 95% confidence intervals (CIs) associated with a 1 standard deviation (SD) increase in continuous variables or compared to the indicated reference group for categorical variables. Model results are presented as unadjusted associations and after adjustment for age, race, and gender, with and without BMI adjustment.

In addition to these unadjusted and adjusted bivariate analyses, we performed multivariate analyses within the subset of the sample with available data on all measures from pictures P1, P2 and P3, which were available on a large proportion of the photographed samples. Given the discovery nature of these analyses, for both OSA status and continuous AHI, the "best model" was determined using a backwards selection algorithm and a p-value threshold of 0.05 for retaining an individual variable in the final model.

We note that the sample size varied across photographs, as not all photographs were taken in every subject and differential visibility of the anatomy caused further variation in sample size across measures (details regarding visibility and the number of missing values are shown in **e-Table 4**, below). The average sample size for a given measurement was approximately 600. Using this average sample and an α=0.05, we had approximately 90% power to detect an odds ratio of 1.35 associated with a 1 standard deviation (SD) increase in our continuous measures or across categorical measures. Similarly, we estimate >90% power to detect a mean difference between cases and controls of 0.28 standard deviations in any given measure.

Given the exploratory nature of analyses, as well as the correlation between measures in similar domains, uncorrected p-values are presented for all analyses, with a $p<0.05$ was considered evidence for a nominally significant association. To control for multiple comparisons, we utilized a Hochberg step-up method [6, 7] for controlling the overall family-wise error rate (α) at 0.05 for a given analysis. Briefly, for a given set of k null hypotheses (H_{0i}) we denote the associated p-values as p_i, where i=1, 2, … k represents the rank of each p-value ordered from smallest to largest. Given this ranking, the Hochberg step-up procedure uses the following procedure to determine which null hypotheses to reject:

Step 1: *If $p_k < \alpha$, reject H_{0i} for i=1, 2, … k; else go to Step 2;*

Step 2: *If $p_{k-1} < \alpha/2$, reject H_{0i} for i=1, 2, … k-1; else go to Step 3;*

[…]

Step k: *If $p_i < \alpha/k$, reject H_{0i} for i=1; else, no null hypotheses are rejected*

Statistically significant p-values are highlighted in bold within relevant tables. All analyses were conducted using STATA, Version 14 (StataCorp LP, College Station, TX) or SAS Version 9.4 (SAS Institute, Cary, NC).

SUPPLEMENTAL RESULTS

The absolute error of the laser ruler was assessed with repeated measures from 40 cm of a circle with diameter of 5.08 cm (i.e., 2 inches) and area of 20.27 cm^2 (**e-Table 1**). The digital morphometrics method slightly underestimated the diameter by 0.01 cm (0.2%) and slightly overestimated the area by 0.26 cm^2 (1.3%). Low coefficients of variation (<1%) for the diameter and area indicate that measurements are accurate and reproducible for repeated measures at the same distance (**e-Table 1**).

The effect of distance on digital morphometrics measurements was examined by repeating photographs every centimeter from distances of 35-50 cm (**e-Table 2**). We observed similar average estimates and low coefficients of variation to those in the pictures repeated at 40 cm (**e-Table 1**). There was no significant correlation between distance and measured values. Thus, measurements were reproducible across distances.

We validated our analysis technique by re-analyzing photographs from 10 randomly selected subjects at least 1 week apart. We observed high reproducibility of continuous measures, with ICCs>0.93 for all measures except tongue thickness, which still had an acceptable ICC of 0.77 (**e-Table 3**). For categorical variables, which require more subjective assessment, we observed good percent agreement (>70% for all measures). Disagreements were within one unit for all measures except pharyngeal narrowing grade.

To evaluate the photographic reproducibility, we took repeat photographs in 10 subjects (convenience sample), within a 1-month interval. Continuous measures had good reproducibility, with all ICCs≥0.76 (**e-Table 3**). Categorical measurements showed acceptable levels of agreement (≥80% for all measures); any disagreements were within one unit.

Finally, the relationship between Mallampati calculated clinically and that from digital pictures is of interest. We note that both measures rely on the same definitions. However, primary benefits of basing the scoring on a digital photograph are (1) a standardized procedure for capturing the inside of the mouth across all individuals and (2) a stored record of the view of the mouth (i.e., the photo) for reference and/or rescoring. While this study was not designed to robustly explore the relationship between our digital assessment of Mallampati and clinical assessment, we extracted 'clinically assessed' Mallampati scores from the medical records of 100 recent patients that were included in our analyses. Interestingly, 94 of these patients had the highest Mallampati score of Class IV, with 1 patient being scored as Class I, 1 as Class II and 4 patients as Class III. While it is difficult to perform robust association tests given the high percentage of Class IV values, we note that there was 80% agreement between clinical and Modified Mallampati scores based on the digital photographs. Differences generally reflect that fact that, while a high proportion of these 100 patients also had Class IV scores using our digital photographs to characterize Modified (83%) Mallampati, there were more scores of Class I, II or III. This may suggest that there is more accurate/specific grading when using a digital photo reference as opposed to scoring Mallampati within a clinic visit; however, we note that the photo scoring was not performed by the same clinician as the values derived from physical examination. Future studies should examine this question more directly, including a wider range of Mallampati scores and 'clinical' and photographic scoring performed by the same individual.

SUPPLEMENTAL DISCUSSION

Limitations and Strengths

We used an AHI threshold of 10 events/hour to distinguish between normal subjects and patients with sleep apnea. This cut off is higher than what has been used for most OSA studies (typically the controls have an AHI <5 events/hour) [8]. However, other studies have also used an AHI of 10 events/hour to define controls [9, 10]. Any classification bias from including patients with AHI between 5-10 as controls, rather than excluding these participants or using an AHI<5 definition is expected to make it more difficult to find a difference between apneics and controls. In future studies we will examine differences in upper airway anatomic changes using digital morphometrics with different AHI severities.

A small proportion of the subjects (8.5%) had home sleep studies, rather than in-lab polysomnography. We believe that the potential for bias due to differences in sleep study methods in our sample is minimal. First, only 73 (8.5%) of the 860 patients in our study sample had home sleep tests. Moreover, recent studies [11-13] comparing HST and PSG suggest that, although there are systematic deviations, the AHI on HST is only ~2.5 events/hour lower on average when compared to PSG values in the same individual. If we inflate the AHI by this amount within the patients that underwent an HST in our study, only 10 (13.7%) patients would have been called cases instead of controls using our definition of AHI≥10 events/hour. Given that these 10 individuals represent only 1.2% of our overall sample, potential misclassification bias is minimal. Finally, we note that this type of misclassification bias would be expected to make it more difficult to find associations with OSA, as it leads to apneics being more like control patients with respect to AHI.

The laser ruler shows minimal variation across distances and over time, however, the procedure slightly overestimated the structural size of an object. The reproducibility and precision of the laser ruler was robust. Nevertheless, it is important to recognize that these calculated measures are likely slight overestimates of the actual structure size, although the clinical relevance of these differences is minimal.

During the validation and reproducibility analyses, in general we observed slightly worse agreement for subjective, categorical measures than more objective structure quantifications; however, all measures met our minimum level for acceptable reproducibility. The lower reproducibility of Mallampati reflects difficulties previously documented with modified Mallampati [14]. For instance, as the subject breathes, the tongue and soft palate move, creating a fluid dynamic between two or more modified Mallampati classes. However, differences in modified Mallampati score were never more than 1 class and differences in pharyngeal airway visibility, a dichotomous version of modified Mallampati, showed 100% agreement. By virtue of including 4 categories, modified Mallampati provided more information regarding upper airway crowding than the dichotomous score of pharyngeal airway visibility. However, pharyngeal airway visibility was more reproducible and consequently may prove more useful than modified Mallampati in describing intraoral crowding as it relates to OSA. Future studies should examine techniques for objectively determining these measures of airway crowdedness.

Limited visibility of the uvula, airway, lateral wall, and tonsils, which was related to lower age, higher obesity, and less likelihood of having OSA, limits the generalizability of results involving these measures. Future studies should seek to develop high-throughput procedures for accurately capturing these structures in an efficient manner through photography in a larger proportion of patients. The ability to quantify this anatomy in a specific subset of the population could be useful in a clinical setting when evaluating risk for sleep apnea. In addition, the inability to visualize the pharyngeal airway could be a phenotype for sleep apnea, as discussed above. Similarly, multivariate analyses were restricted to patients with data available for all measures; while we restricted these analyses to the subset of measures available in a high proportion of the targeted sample, differences in the number of patients with each picture and missing values led to a final sample of 430 patients. While this reduced sample could limit generalizability of these analyses, we found no differences in clinical covariates (age, gender, race, BMI or AHI) between included and excluded individuals. Overall, while measurements of some structures was limited, our protocol was able to obtain tongue and mouth measurements in almost all subjects, indicating that digital morphometrics with a laser ruler can be utilized in apneics to routinely obtain these characteristics.

Any study investigating the physiology of OSA is remiss to ignore measures of craniofacial anatomy.

The goal of the current study was to describe a novel technique for quantifying intraoral measurements of the tongue, mouth and pharyngeal crowdedness; the laser ruler technique could be used to perform comprehensive assessments of craniofacial structures via similar techniques to those presented previously by Lee *et al*.[15] In order to comprehensively address the pathogenesis of OSA in future analyses, the intraoral and craniofacial measures must be examined together to develop a complete morphometric model.

There are several important strengths to this proposal: 1) we had a very large sample of controls and apneics; 2) this is the first study to use a laser ruler to quantify intraoral structures; 3) the morphometric technique is reproducible, reliable, does not involve any radiation and can be performed quickly; 4) we were able to quantify differences in pharyngeal anatomy between controls and apneics.

e-Table 1: Absolute Error of the Laser Device based on 10 Measures of a Circle with 2 Inch Diameter

	Diameter (cm)	Area (cm^2)
Actual value	5.08	20.27
Measured Mean ± SD	5.07 ± 0.02	20.53 ± 0.17
Difference (Mean-Actual)	-0.01	0.26
Coefficient of Variation[†]	0.41%	0.82%
Individual Trials[‡]		
1	5.08	20.58
2	5.08	20.62
3	5.05	20.39
4	5.09	20.77
5	5.05	20.35
6	5.05	20.37
7	5.11	20.79
8	5.08	20.64
9	5.07	20.46
10	5.05	20.38

[†]Coefficient of Variation = [SD/Mean]*100%; [‡]A series of 10 photographs were taken of a circle with a 2 inch diameter from a distance of 40 cm and the diameter and area were computed.

e-Table 2: Effect of Distance based on Measures of a Circle with 2 Inch Diameter

	Diameter (cm)	Area (cm^2)
Actual value	5.08	20.27
Measured Mean ± SD	5.11 ± 0.03	20.73 ± 0.20
Difference (Mean-Actual)	0.03	0.46
Coefficient of Variation[†]	0.58%	0.95%
Individual Trials[‡,§]		
35cm	5.10	20.65
36cm	5.11	20.75
37cm	5.14	21.01
38cm	5.16	21.06
39cm	5.07	20.57
40cm	5.08	20.58
41cm	5.07	20.47
42cm	5.11	20.85
43cm	5.12	20.86
44cm	5.11	20.72
45cm	5.06	20.52
46cm	5.08	20.49
47cm	5.15	21.09
48cm	5.10	20.66
49cm	5.12	20.72
50cm	5.11	20.77

[†]Coefficient of Variation = [SD/Mean]*100%; [‡]One photograph was taken of a circle with 2 inch diameter in 1 cm increments from 35 cm to 50 cm and the diameter and area were computed; [§]No significant correlation between distance and either diameter (-0.04, p=0.894) or area (-0.08, p=0.766).

CHEST® Online Supplement

e-Table 3: Reproducibility of Photography and Analysis Validation

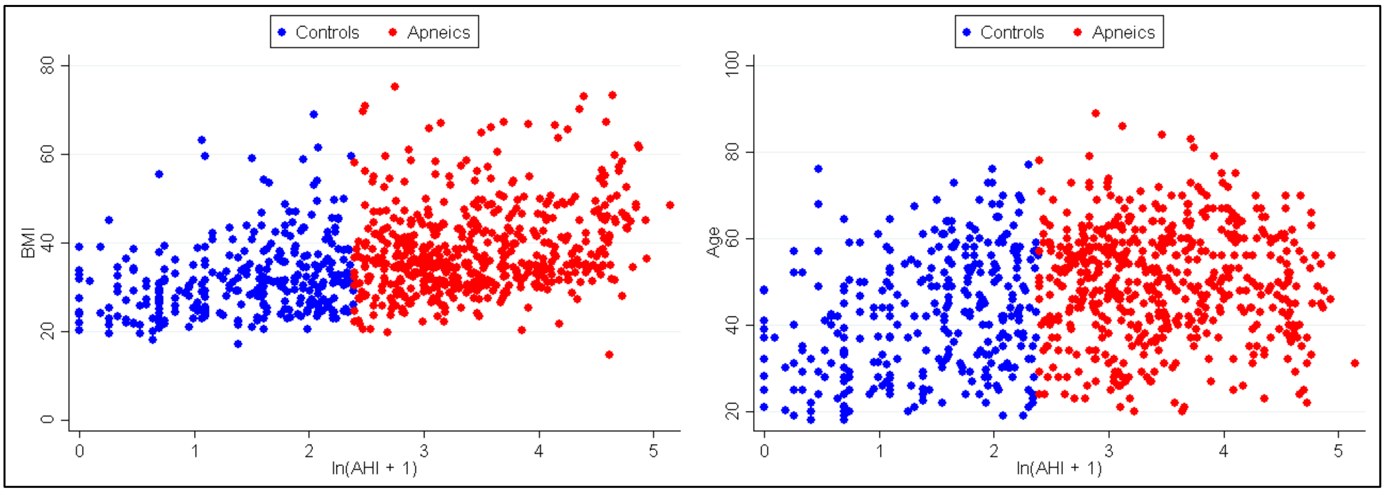
Picture	Measure	Analysis Validation		Photographic Validation	
		N Pairs	ICC or % Agreement	N Pairs	ICC or % Agreement
P1	Modified Mallampati	10	80%	10	90%
	Airway Visibility	10	100%	10	100%
	Mouth Width	10	0.969	10	0.776
	Mouth Height	10	0.998	10	0.849
	Mouth Area	10	0.999	10	0.882
	Tongue Width	10	0.948	10	0.907
P2	Standard Mallampati	10	90%	10	80%
	Mouth Width	10	0.987	10	0.850
	Tongue Width	10	0.992	10	0.895
	Tongue Length	10	0.942	10	0.764
	Tongue Area	10	0.970	10	0.890
P3	Tongue Length	10	0.992	10	0.877
	Tongue Area	10	0.935	10	0.939
	Tongue Thickness	10	0.768	10	0.768
	Tongue Curvature	10	0.931	10	0.940
P4	Airway Width	10	0.974	10	0.862
	Uvula Length (Airway)	10	0.952	10	0.801
	Uvula Width (Airway)	10	0.953	10	0.843
	Uvula Area (Airway)	10	0.951	10	0.787
Global	Mallampati	10	100%	10	80%
	Evident Tongue Ridging	10	90%	10	80%
	Tonsil Hypertrophy Grade	6	100%	7	100%
	Pharyngeal Narrowing Grade	7	71%	10	100%

e-Table 4: Visibility and Measurability of Continuous Intraoral Measurements Requiring Lasers

Picture	Total Pictures	Measure	Total Visible	Controls (AHI<10)	Cases (AHI≥10)	p
P1	816	Mouth Width	808 (99.0%)	308 (99.4%)	500 (98.8%)	0.717
		Mouth Height	780 (95.6%)	298 (96.1%)	482 (95.3%)	0.556
		Mouth Area	786 (96.3%)	301 (97.1%)	485 (95.9%)	0.358
		Tongue Width	742 (90.9%)	**290 (93.6%)**	**452 (89.3%)**	**0.042**
P2	812	Mouth Width	725 (89.2%)	282 (91.3%)	443 (88.1%)	0.154
		Tongue Width	798 (98.3%)	306 (99.0%)	492 (97.8%)	0.196
		Tongue Length	740 (91.1%)	277 (89.6%)	463 (92.1%)	0.242
		Tongue Area	750 (92.4%)	281 (90.9%)	469 (93.2%)	0.230
P3	642	Tongue Length	624 (97.2%)	245 (97.6%)	379 (96.9%)	0.611
		Tongue Area	622 (96.9%)	243 (96.8%)	379 (96.9%)	0.933
		Tongue Thickness	626 (97.5%)	246 (98.0%)	380 (97.2%)	0.515
		Tongue Curvature	615 (95.8%)	242 (96.4%)	373 (95.4%)	0.531
P4	771	Airway Width	148 (19.2%)	**88 (29.7%)**	**60 (12.6%)**	**<0.0001**
		Uvula Length (Airway)	169 (21.9%)	**91 (30.7%)**	**78 (16.4%)**	**<0.0001**
		Uvula Width (Airway)	358 (46.4%)	**165 (55.7%)**	**193 (40.6%)**	**<0.0001**
		Uvula Area (Airway)	166 (21.5%)	**89 (30.1%)**	**77 (16.2%)**	**<0.0001**

Significant differences between cases and controls are shown in **bold**.

e-Figure 1: *Overlap between BMI, Age and Natural Log Transformed AHI.* The relationship between OSA severity, measured as ln(AHI+1) and age and BMI are illustrated in controls (blue dots) and apneics (red dots). There is the expected positive correlation between these measures and disease severity. Note the overall overlap in the distributions, which allows for appropriate adjustment in the statistical models.



SUPPLEMENTAL REFERENCES

1. Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, Marcus CL, Mehra R, Parthasarathy S, Quan SF, Redline S, Strohl KP, Davidson Ward SL, Tangredi MM. Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 2012: 8(5): 597-619.

2. Barbera AL, Sampson WJ, Townsend GC. An evaluation of head position and craniofacial reference line variation. *Homo* 2009: 60(1): 1-28.

3. Weiss TM, Atanasov S, Calhoun KH. The association of tongue scalloping with obstructive sleep apnea and related sleep pathology. *Otolaryngol Head Neck Surg* 2005: 133(6): 966-971.

4. Friedman M, Ibrahim H, Joseph NJ. Staging of obstructive sleep apnea/hypopnea syndrome: a guide to appropriate treatment. *Laryngoscope* 2004: 114(3): 454-459.

5. Brodsky L. Modern assessment of tonsils and adenoids. *Pediatr Clin North Am* 1989: 36(6): 1551-1569.

6. HOCHBERG Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 1988: 75(4): 800-802.

7. Huang Y, Hsu JC. Hochberg's Step-Up Method: Cutting Corners Off Holm's Step-Down Method. *Biometrika* 2007: 94(4): 965-975.

8. Kushida CA, Littner MR, Morgenthaler T, Alessi CA, Bailey D, Coleman J, Jr., Friedman L, Hirshkowitz M, Kapen S, Kramer M, Lee-Chiong T, Loube DL, Owens J, Pancer JP, Wise M. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep* 2005: 28(4): 499-521.

9. Kim AM, Keenan BT, Jackson N, Chan EL, Staley B, Poptani H, Torigian DA, Pack AI, Schwab RJ. Tongue fat and its relationship to obstructive sleep apnea. *Sleep* 2014: 37(10): 1639-1648.

10. Kim AM, Keenan BT, Jackson N, Chan EL, Staley B, Torigian DA, Alavi A, Schwab RJ. Metabolic activity of the tongue in obstructive sleep apnea. A novel application of FDG positron emission tomography imaging. *Am J Respir Crit Care Med* 2014: 189(11): 1416-1425.

11. Cairns A, Wickwire E, Schaefer E, Nyanjom D. A pilot validation study for the NOX T3(TM) portable monitor for the detection of OSA. *Sleep Breath* 2014: 18(3): 609-614.

12. Ng SS, Chan TO, To KW, Ngai J, Tung A, Ko FW, Hui DS. Validation of Embletta portable diagnostic system for identifying patients with suspected obstructive sleep apnoea syndrome (OSAS). *Respirology* 2010: 15(2): 336-342.

13. Ayappa I, Norman RG, Seelall V, Rapoport DM. Validation of a self-applied unattended monitor for sleep disordered breathing. *J Clin Sleep Med* 2008: 4(1): 26-37.

14. Hukins C. Mallampati class is not useful in the clinical assessment of sleep clinic patients. *J Clin Sleep Med* 2010: 6(6): 545-549.

15. Lee RW, Chan AS, Grunstein RR, Cistulli PA. Craniofacial phenotyping in obstructive sleep apnea--a novel quantitative photographic approach. *Sleep* 2009: 32(1): 37-45.