Association of Palatine Tonsil Size and Obstructive Sleep Apnea in Adults

Sebastian M. Jara, MD[1] and **Edward M. Weaver, MD, MPH**[1,2]

[1]Department of Otolaryngology – Head and Neck Surgery, University of Washington, Seattle, WA

[2]Department of Otolaryngology – Head and Neck Surgery, Veterans Affairs Puget Sound Health Care System, Seattle, WA

Abstract

OBJECTIVES—The relationship between palatine tonsil (PT) size and obstructive sleep apnea (OSA) has not been well established in adults. The purpose of this study is to test the association between PT grade, PT volume, and OSA severity in U.S. adult patients.

STUDY DESIGN—Cross-sectional study of all patients (age 18 years) who underwent pharyngeal surgery for OSA that included palatine tonsillectomy with tonsil volume measurement, from January 2011 to June 2016.

METHODS—Medical records were reviewed for PT grade (measured on clinical exam by the Brodsky tonsil grading scale), PT volume (measured intra-operatively by water displacement), and apnea-hypopnea index (AHI). Associations were evaluated with multivariate linear regression adjusting for age, sex, body mass index, smoking status, lingual tonsil volume (AHI models only), and multilevel surgery aside from lingual tonsillectomy (PT volume versus AHI model only).

RESULTS—The cohort (N=83) was middle-aged (mean age 43+/−12 years), predominantly male (61%), obese (mean body mass index 33+/−7 kg/m^2), and had severe OSA (mean AHI 32+/−28). After adjustment for confounders, PT grade was strongly associated with PT volume (Beta=1.8, 95%CI: [1.0,2.6], P<0.001) and with AHI (Beta=13.5, 95%CI: [3.5,23.6], P=0.01); PT volume was not associated with AHI (Beta=−0.2, 95%CI: [−2.2,1.9], P=0.89).

CONCLUSIONS—In contrast to past studies, subjective PT grade (versus objective PT volume) was more strongly associated with AHI. These data suggest that the space the tonsils occupy within the oropharyngeal airway, instead of their actual measured volume, may be more predictive of OSA severity in a cohort of U.S. adult patients.

Keywords

obstructive sleep apnea; surgical treatment of obstructive sleep apnea; clinical/outcomes research

---

Address for correspondence: Sebastian M. Jara, MD, Department of Otolaryngology – Head and Neck Surgery, University of Washington, 1959 NE Pacific St, Box 356515, Seattle, WA 98195-6515, Phone: (617) 791-0506, Fax: (206) 543-5152, sebastian.jara@gmail.com.

INTRODUCTION

Obstructive sleep apnea (OSA) is a disorder characterized by symptomatic, repeated upper airway obstruction during sleep. The source of obstruction is frequently multifactorial and may involve several anatomic sites in the upper airway. Palatine tonsil (PT) hypertrophy represents a frequently encountered potential source of upper airway obstruction. Nonetheless, the relationship between PT size and degree of OSA severity remains unclear.

In children with OSA, tonsillar hypertrophy is the most frequent source of obstruction.[1] However, it has not been established that increasing tonsil size necessarily correlates to worsening OSA severity. A systematic review assessing tonsil size and pediatric OSA severity found a weak positive association between tonsil size and OSA severity in half of the studies assessed, with higher-quality studies showing no association.[2] Similarly, a recent study found no correlation between PT grade and OSA severity.[3] Another recent study showed that smaller PT grade demonstrates less oropharyngeal obstruction during drug-induced sleep endoscopy, suggesting other possible sources of obstruction in children with OSA and small tonsils.[4]

Few of studies in adults have investigated the association between PT grade, PT volume, and OSA severity.[5–7] Wang et al. demonstrated a positive correlation between PT grade with PT volume in a Korean cohort of patients.[5] Similarly, Cahali et al. showed a positive correlation between PT volume and AHI, adjusted for covariates, in a Brazilian cohort of OSA patients.[6] More recently, Lai et al. demonstrated a significant unadjusted correlation between PT volume and AHI, but weak correlation between PT grade and AHI, in a Taiwanese cohort of OSA patients.[7] While these associations have been tested in homogeneous patient cohorts, they have not been tested in more heterogeneous patient cohorts, like represented in U.S. adult patients, through adjusted analyses.

The purpose of this study was to test the independent association between subjective and objective measures of PT size and OSA severity in a heterogeneous cohort of U.S. adult patients.

METHODS

Study Design and Subjects

This was a cross-sectional study of all adult patients (age ≥ 18 years) who underwent pharyngeal surgery for the management of OSA that included palatine tonsillectomy and intraoperative measurement of tonsil volume. Patients treated by a single sleep surgeon from January 1, 2011 to June 30, 2016 were identified. All patients had clinic notes, operative reports, and pre-operative polysomnography reports available for review. Subjective palatine tonsil (PT) grade was measured during pre-operative clinical evaluation by a single provider and graded using the Brodsky tonsil grading scale as follows: 0, surgically absent; 1+, 0–25% oropharyngeal obstruction; 2+, 25–50% obstruction; 3+, 50–75% obstruction; and 4+, 75–100% obstruction.[8] For instances in which patients exhibited tonsillar asymmetry on evaluation (i.e different grades documented for each tonsil), patients were assigned an overall PT grade corresponding to the higher grade of the two tonsils for subsequent

analyses, as this best approximated degree of oropharyngeal obstruction noted on exam. Objective PT volume was measured during surgery by volumetric water displacement of the excised tonsils in a graduated cylinder. If a lingual tonsillectomy or partial glossectomy was performed, the volume of each respective surgical specimen was measured separately using water displacement. Baseline OSA severity measures (e.g. apnea-hypopnea index [AHI]) were extracted from pre-operative polysomnography reports. Medical records were reviewed for additional clinical data including baseline demographics, body mass index (kg/m^2), current smoking status (yes/no), and pre-operative Epworth Sleepiness Scale. Institutional Review Board approval was obtained from the University of Washington prior to data collection.

Statistical Analysis

All statistical analyses were conducted using Stata 14.2 (Stata Inc., College Station, TX). Baseline patient characteristics were summarized using means, standard deviations, and frequency. Spearman's rank correlation was used to test the univariate correlation between PT grade and PT volume, PT grade and AHI, and PT volume and AHI, respectively. Multivariate linear regression models were created to further test these associations adjusting for age, sex, body mass index, lingual tonsil volume (AHI models only) and multilevel surgery besides lingual tonsillectomy (PT volume versus AHI model only). Lingual tonsil volume was used to adjust the AHI models (i.e. PT grade versus AHI and PT volume versus AHI) as lingual tonsil volume is potentially associated with PT size (measured subjectively or objectively) and with AHI, and thus is a potential confounder for this association. Similarly, the need for multilevel surgery aside from lingual tonsillectomy (defined as surgery including partial glossectomy, direct epiglottis procedures, genioglossus advancement, or hyoid suspension, in addition to uvulopalatopharyngoplasty) was viewed as a potential confounder for the association between PT volume and AHI, and correspondingly used to adjust this multivariate model. Nasal surgery was not used as a level in the designation of multilevel surgery. Secondary multivariate analyses were performed using additional OSA severity and symptom metrics (i.e. apnea index, oxygen desaturation index, lowest O_2 saturation, and Epworth Sleepiness score). For all tests, $P < 0.05$ was considered statistically significant.

RESULTS

The cohort consisted of 83 adult patients who underwent palatine tonsillectomy (Table 1). The cohort was middle-aged, mostly male, mostly White, and obese, on average (Table 1). Most patients had PT grade 1+ to 3+ and AHI generally increased with increasing PT grade (Table 1). Two patients were noted to have tonsillar asymmetry (1+/2+ and 2+/3+) and were categorized according to their higher PT grade (2+ and 3+, respectively). Interestingly, one patient had undergone previous tonsillectomy and was clinically grade 0 during pre-operative evaluation, but had residual palatine tonsil tissue identified and removed during surgery. Patients had excessive daytime sleepiness and severe OSA on average (Table 1). Most patients had concurrent uvulopalatopharyngoplasty and many had other procedures (Table 2). Only one patient (1%) underwent palatine tonsillectomy alone.

Bivariate analysis with Spearman's rank correlation and simple linear regression demonstrated statistically significant positive correlations between PT grade and PT volume (r = 0.64, P < 0.001) and PT grade and AHI (r = 0.22, P = 0.04). The correlation was small and insignificant between PT volume and AHI (r = 0.11, P = 0.36). Correlations tested using OSA severity (i.e. mild, moderate, and severe) were comparable to those obtained from correlations using AHI as a continuous variable (data not shown). Figure 1 shows scatterplots between PT grade and PT volume, PT grade and AHI, and PT volume and AHI and Table 3 shows the simple (unadjusted) regressions. Results were similar after adjustment for confounding variables (Table 3). Similarly, PT grade was independently associated with other objective OSA measures, and PT volume was not associated with other OSA measures (Tables 4 & 5). Neither PT grade nor PT volume were independently associated with sleepiness (Tables 4 & 5).

DISCUSSION

In this U.S. adult cohort, we observed a positive correlation between subjective PT grade and objective PT volume. According to the beta coefficient, every increase in PT grade (0 to +1, +1 to +2, etc.), was associated with an average increase in PT volume of ~2 mL. This correlation was observed in both the unadjusted model and after adjusting for age, sex, body mass index, and smoking. Previous studies in both adult and pediatric patients with OSA have noted similar findings.[5–7,9] Lai et al.[7] found statistically significant associations between PT grade and PT volume in unadjusted analyses. Their group also observed significant correlations between PT grade and other objective measures of PT size (i.e. length and weight). Cahali et al.[6] also found a positive correlation between PT grade and PT volume, adjusted for age, body mass index, Mallampati grade, and Friedman stage. Our results are consistent with past findings and suggest that subjectively-assessed PT grade is a predictor of objectively measured PT volume in adults.

We found that increasing PT grade was significantly associated with increasing OSA severity. The magnitude of this association - as denoted by the beta coefficient (and associated 95% CI) between PT grade and AHI – became greater in the multivariate analysis after adjustment for age, sex, body mass index, smoking, and lingual tonsil volume. With an adjusted beta coefficient of almost 14, these data indicate that every increase in PT grade was associated with an average increase in AHI of ~14 events/hour (compared ~9 events/ hour in the unadjusted model). Similarly, PT grade had comparable association with other polysomnography measures.

In contrast, we found no association between PT volume and OSA severity, whether measured by AHI or polysomnography parameters, even after adjustment for potential confounding variables. Unexpectedly, these results suggest that subjective, pre-operative PT grading is more strongly associated with OSA severity than objective, intraoperative measurement of PT volume in a U.S. cohort.

These latter results differ significantly from past studies. Lai et al.[7] found a significant correlation between objective PT measures (volume and weight) and OSA severity (AHI and snoring index). However, they found only a borderline correlation between PT grade and

OSA severity, and concluded that objective tonsil measures were more meaningful for predicting OSA severity than subjective tonsil grading. Similarly, Cahali et al.[6] reported significant correlations between increasing PT volume and increasing AHI as well as increasing Epworth Sleepiness score, but did not observe this correlation with increasing PT grade. Past studies in pediatric patients have similarly shown correlations between increasing PT volume, but not grade, and OSA severity.[3,9]

These discrepancies may be due to the greater variability in patient cohort features in our cohort compared to the other cohorts. Both Lai et al. and Cahali et al. studied more homogenous patient cohorts (Taiwanese and Brazilian patient populations, respectively) while the current study was performed in a heterogeneous cohort of U.S. patients. It is well-established that the craniofacial bony skeleton contributes to the anatomic pathogenesis of OSA, with variants in bone structure and degree of bony development contributing to the volume of the adult upper airway.[10] Thus, it is possible that PT volume (compared to PT grade) may be a more significant predictor of OSA severity in populations with more homogenous and constricted craniofacial features, as seen in homogeneous Asian population and in children. In contrast, in a U.S. cohort of adult patients with more heterogeneous craniofacial features, and thus greater ranges of anatomic upper airway size, the actual volume of the tonsils may matter less than the proportion of space they occupy within the oropharyngeal airway, which is better assessed by PT grade.

In addition, the patient cohort in this current study had a greater variability in weight. The patients in this cohort were obese on average and exhibited considerable variability in body mass index, ranging from normal weight to morbid obesity (Table 1). In contrast, Lai et al. and Cahali et al had patient cohorts with both lower and less variable body weights (body mass indices 27 +/− 3 kg/m^2 and 29 +/− 3 kg/m^2, respectively).[6,7] In fact, Cahali et al. excluded patients with a body mass index > 35 kg/m^2 to reduce potential variability in oropharyngeal features.[6] As such, the greater variability in weight, and subsequent effects on tongue base volume[11], may have contributed to greater variability in upper airway volume in our cohort, making PT grade a more significant predictor of OSA severity than PT volume.

Lastly, neither of the prior studies adjusted their analyses of PT grade and OSA severity for potential confounders. Because several factors are associated with both PT size and OSA severity – including age, sex, body mass index, and smoking – it is possible that these factors confounded the measured association between PT size and OSA severity. As such, because our study assessed a patient cohort with greater variability with these features, the results of our study may be more generalizable and broadly applicable to adult OSA patients.

There are several limitations to this study. One limitation was the diversity of obstructive anatomy present in our cohort. Although a range of PT sizes were well-represented, our cohort had a diverse range other obstructive anatomy, with many patients undergoing multilevel surgery. This is similar to the study by Lai et al, in which patients also underwent multilevel surgery, but contrasts the study by Cahali et al, in which all patients underwent single level surgery (tonsillectomy as component of either uvulopalatopharyngoplasty or lateral pharyngoplasty).[6,7] Thus, the role of tonsillar enlargement may have had a greater

obstructive influence in our patients, especially when measured categorically by PT grade, than in the patients assessed by Cahali et al, providing an additional potential source of discrepancy in the results we observed. Additionally, because the presence of multilevel obstruction may have influenced measures of OSA severity, and thus our measured associations, this may represent a source of bias in this study. We attempted to adjust for this potential confounding, albeit imperfectly, with lingual tonsil volume and concurrent multilevel surgery, and the results did not change. While this potential confounding is a limitation, it is not clear that it presents a systematic bias to explain the difference between cohorts.

A second limitation is that we were not able to test the prognostic ability of PT grade or PT volume on surgical outcome, because almost every case had multiple procedures. Because many of the patients in this cohort underwent multilevel surgery and multilevel surgery is associated with improved OSA outcomes[12,13], we were unable to correlate PT size and effect of tonsillectomy on postoperative OSA severity. While it has been shown that tonsillectomy improves OSA in adult patients with large tonsils, the effect in adults with small tonsils is less clear.[14–17] Future, prospective studies restricted to patients undergoing isolated tonsillectomy would be needed to test the prognostic ability of PT grade versus PT volume on tonsillectomy outcome. This question was beyond the scope of the current study.

A third limitation is that all clinical and operative findings were assessed by a single sleep surgeon without blinding. Thus, it is theoretically possible that lack of blinding contributed to bias of the measured associations. However, given that data were prospectively collected (even if retrospectively reviewed) prior to formulating the hypothesized associations, this is an unlikely source of bias. Furthermore, because measures were all assessed by a single provider, there was greater uniformity in clinical PT grading as well objective PT volume measurement, and thus is a potential strength of the current study.

CONCLUSION

We found a strong correlation between subjective and objective tonsil measures consistent with the findings of previous studies. However, in contrast to previous studies, we found that subjectively measured PT grade by the Brodsky grading scale was more strongly associated than objectively measured PT volume with OSA severity. Because PT grade is a measure of the fractional oropharyngeal space occupied by the tonsils, these data suggest that PT grade, rather than PT volume, may be more predictive of OSA severity in a heterogeneous cohort of patients with greater upper airway variability.

Acknowledgments

References

1. Baugh RF, Archer SM, Mitchell RB, et al. Clinical practice guideline: tonsillectomy in children. Otolaryngol Head Neck Surg. 2011; 144(1 Suppl):S1–30. [PubMed: 21493257]

2. Nolan J, Brietzke SE. Systematic review of pediatric tonsil size and polysomnogram-measured obstructive sleep apnea severity. Otolaryngol Head Neck Surg. 2011; 144(6):844–850. [PubMed: 21493309]

3. Tang A, Benke JR, Cohen AP, Ishman SL. Influence of Tonsillar Size on OSA Improvement in Children Undergoing Adenotonsillectomy. Otolaryngol Neck Surg. 2015; 153(2):281–285.

4. Miller C, Purcell PL, Dahl JP, et al. Clinically small tonsils are typically not obstructive in children during drug-induced sleep endoscopy. Laryngoscope. Dec.2016 doi: 10.1002/lary.26447

5. Wang JH, Chung Y-S, Jang YJ, Lee B-J. Palatine tonsil size and its correlation with subjective tonsil size in patients with sleep-disordered breathing. Otolaryngol Head Neck Surg. 2009; 141(6):716–721. [PubMed: 19932843]

6. Cahali MB, Soares CF de P, Dantas DA da S, Formigoni GGS. Tonsil volume, tonsil grade and obstructive sleep apnea: is there any meaningful correlation? Clinics (Sao Paulo). 2011; 66(8): 1347–1352. [PubMed: 21915482]

7. Lai C-C, Friedman M, Lin H-C, et al. Objective versus subjective measurements of palatine tonsil size in adult patients with obstructive sleep apnea/hypopnea syndrome. Eur Arch Otorhinolaryngol. 2014; 271(8):2305–2310. [PubMed: 24566883]

8. Brodsky L, Moore L, Stanievich JF. A comparison of tonsillar size and oropharyngeal dimensions in children with obstructive adenotonsillar hypertrophy. Int J Pediatr Otorhinolaryngol. 1987; 13(2): 149–156. [PubMed: 3667094]

9. Howard NS, Brietzke SE. Pediatric tonsil size: Objective vs subjective measurements correlated to overnight polysomnogram. Otolaryngol - Head Neck Surg. 2009; 140(5):675–681. [PubMed: 19393410]

10. Ryan CM, Bradley TD. Pathogenesis of obstructive sleep apnea. J Appl Physiol. 2005; 99(6): 2440–2450. [PubMed: 16288102]

11. Kim AM, Keenan BT, Jackson N, et al. Tongue fat and its relationship to obstructive sleep apnea. Sleep. 2014; 37(10):1639–1648. [PubMed: 25197815]

12. Lin H-C, Friedman M, Chang H-W, Gurpinar B. The efficacy of multilevel surgery of the upper airway in adults with obstructive sleep apnea/hypopnea syndrome. Laryngoscope. 2008; 118(5): 902–908. [PubMed: 18300704]

13. Friedman M, Lin H-C, Gurpinar B, Joseph NJ. Minimally invasive single-stage multilevel treatment for obstructive sleep apnea/hypopnea syndrome. Laryngoscope. 2007; 117(10):1859–1863. [PubMed: 17713449]

14. Smith MM, Peterson E, Yaremchuk KL. The Role of Tonsillectomy in Adults with Tonsillar Hypertrophy and Obstructive Sleep Apnea. Otolaryngol Neck Surg. Mar.2017 19459981769867.

15. Holmlund T, Franklin KA, Levring Jäghagen E, et al. Tonsillectomy in adults with obstructive sleep apnea. Laryngoscope. 2016; 126(12):2859–2862. [PubMed: 27107408]

16. Senchak AJ, McKinlay AJ, Acevedo J, et al. The effect of tonsillectomy alone in adult obstructive sleep apnea. Otolaryngol Head Neck Surg. 2015; 152(5):969–973. [PubMed: 25820584]

17. Camacho M, Li D, Kawai M, et al. Tonsillectomy for adult obstructive sleep apnea: A systematic review and meta-analysis. Laryngoscope. 2016; 126(9):2176–2186. [PubMed: 27005314]
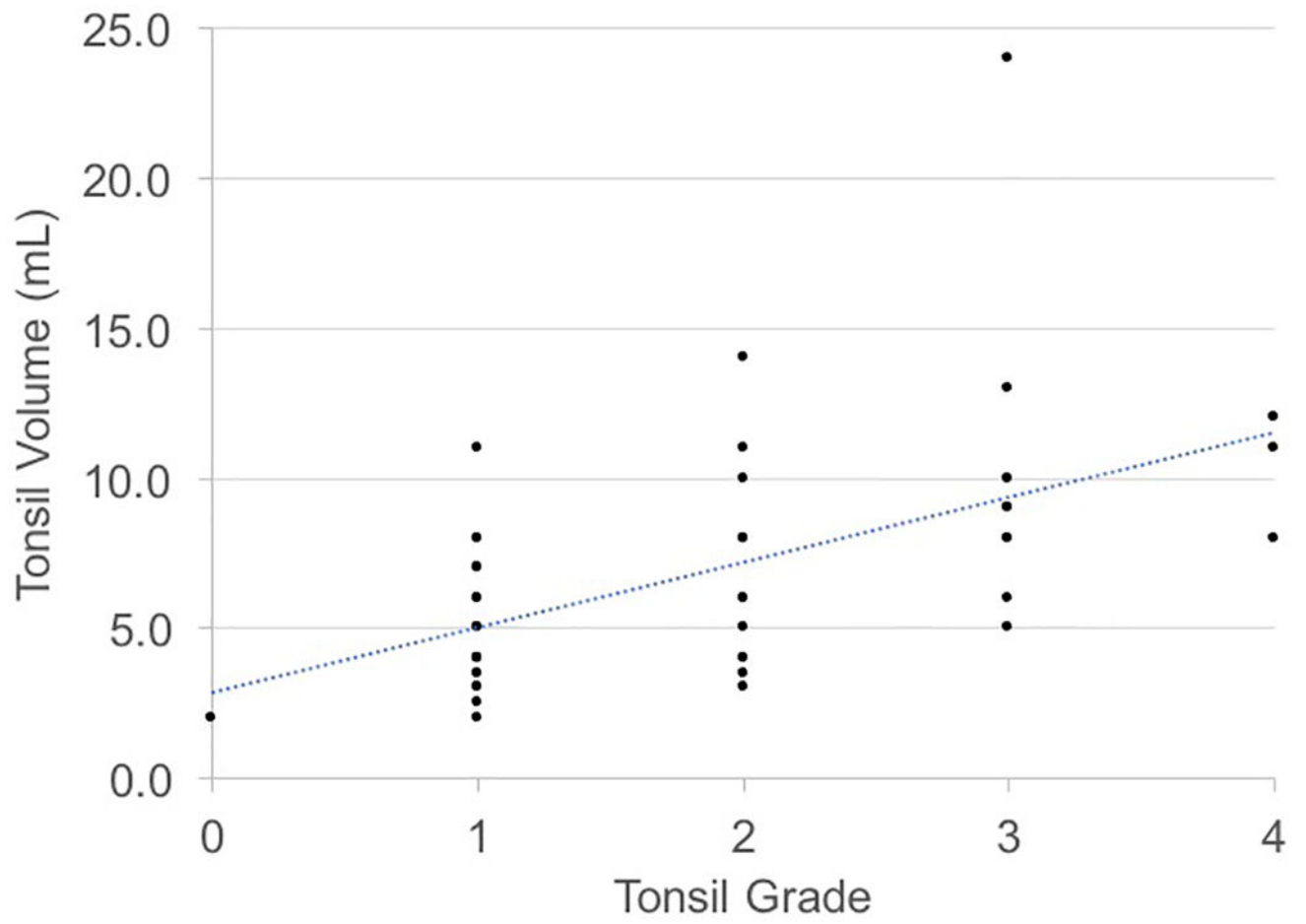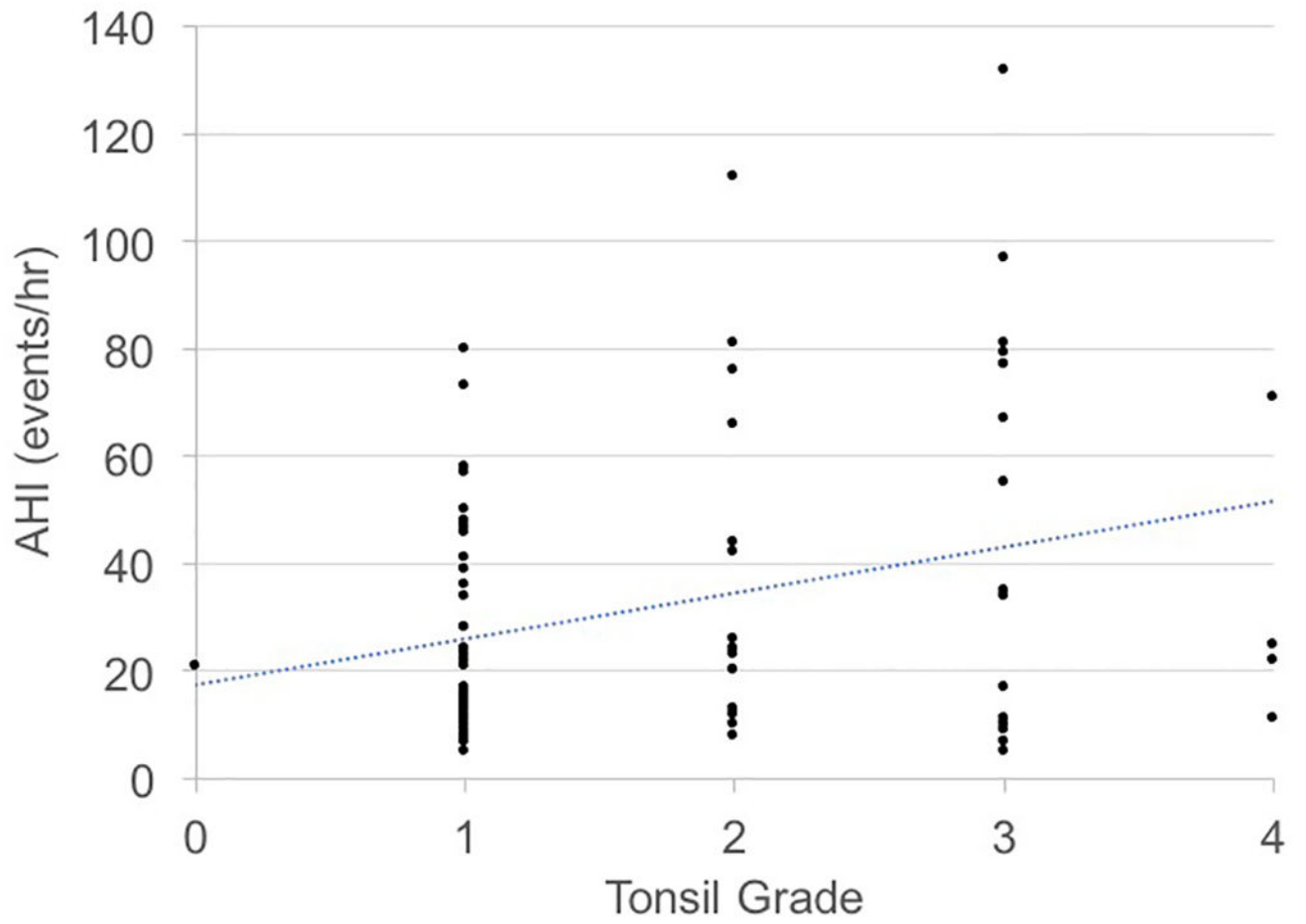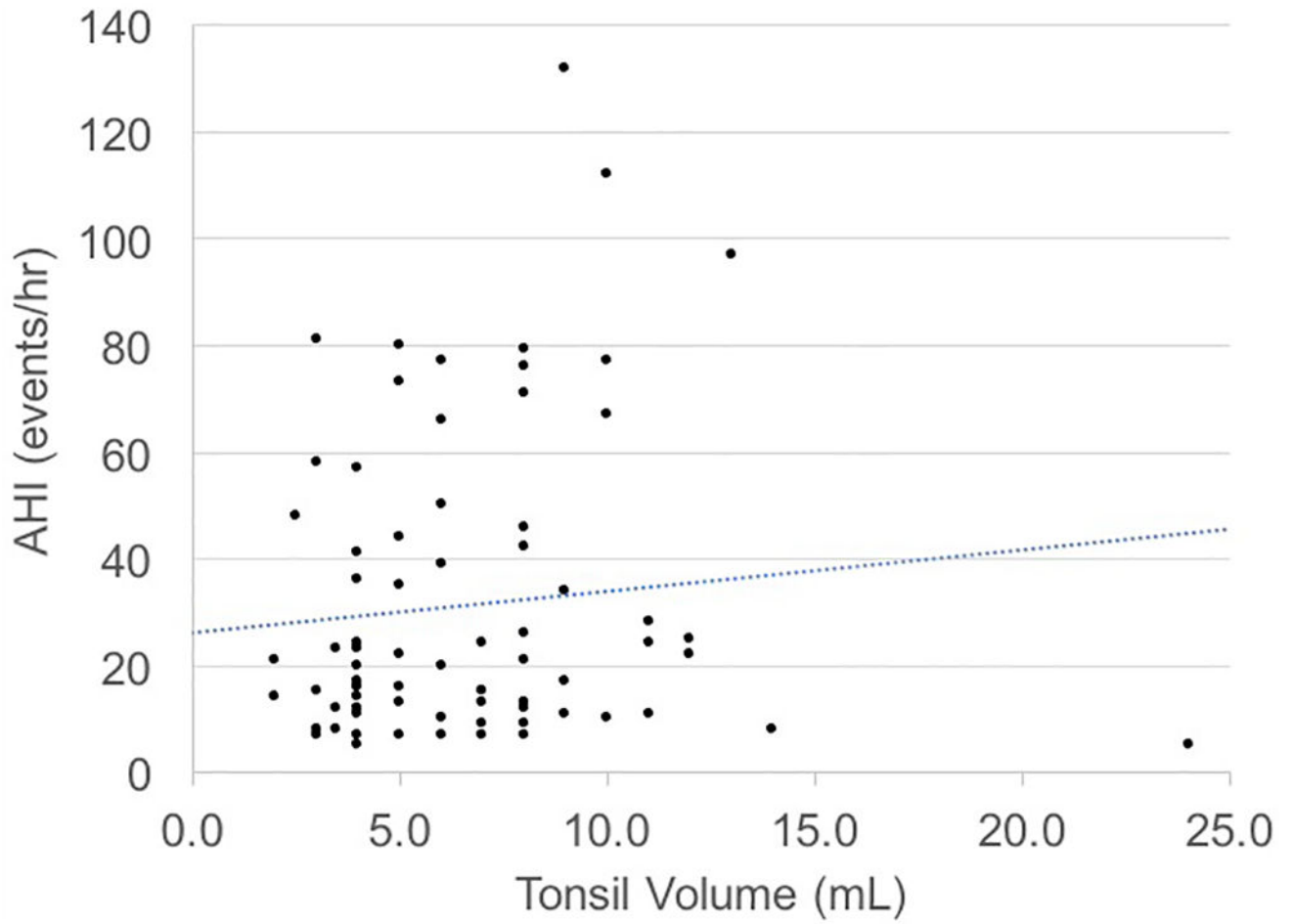
Figure 1.
Relationships between tonsil grade, tonsil volume, and apnea-hypopnea index (AHI).
A) Tonsil Grade vs. Tonsil Volume
B) Tonsil Grade vs. AHI
C) Tonsil Volume vs. AHI

Table 1

Patient Characteristics (N = 83)

Characteristic	Mean +/− SD or %
Age (years)	43 +/− 12
Male (%)	61
White Race (%)	71
Body Mass Index (kg/m2)	33 +/− 7
Current Smoker (%)	11
Palatine Tonsil Grade Distribution (%) and Apnea-Hypopnea Index (events/hr):	
0 (1%)	21[*]
1+ (54%)	25 +/− 3
2+ (21%)	36 +/− 7
3+ (19%)	50 +/− 10
4+ (5%)	32 +/− 13
Palatine Tonsil Volume (mL)	7 +/− 3
Epworth Sleepiness Scale (0–24)	13 +/− 6
Polysomnography Measures:	
Apnea-Hypopnea Index (events/hr)	32 +/− 28
Apnea Index (events/hr)	15 +/− 23
Oxyhemoglobin Desaturation Index (events/hr)	25 +/− 30
Oxyhemoglobin Nadir (%)	82 +/− 9
OSA Severity Distribution (%)	
Mild (AHI 5–15)	39%
Moderate (AHI 16–30)	25%
Severe (AHI > 30)	36%

[*] Data from 1 patient

Table 2

Frequency of Surgeries Performed Concurrently with Palatine Tonsillectomy

Surgical Procedure	Frequency (%)
Uvulopalatopharyngoplasty	96
Lingual tonsillectomy	69
Partial glossectomy	40
Nasal surgery[*]	17
Direct epiglottis procedures[†]	6
Genioglossus advancement	4
Adenoidectomy	4
Hyoid suspension	2

[*] Septoplasty and/or inferior turbinate reduction.

[†] Epiglottoplasty or epiglottopexy.

Table 3

Associations Between Tonsils Sizes and Apnea-Hypopnea Index

Models	Beta Coefficient	95% CI	P-value
Tonsil Grade vs. Tonsil Volume			
Unadjusted model	2.2	(1.5, 2.8)	<0.001
Adjusted model [*]	1.8	(1.0, 2.6)	<0.001
Tonsil Grade vs. Apnea-Hypopnea Index			
Unadjusted model	8.5	(2.4, 14.7)	0.007
Adjusted model [†]	13.5	(3.5, 23.6)	0.01
Tonsil Volume vs. Apnea-Hypopnea Index			
Unadjusted model	0.8	(−1.1, 2.7)	0.42
Adjusted model[‡]	−0.2	(−2.2, 1.9)	0.89

CI = Confidence Interval.

[*] Adjusted for age, sex, body mass index, and smoking status.

[†] Adjusted for age, sex, body mass index, smoking status, and lingual tonsil volume.

[‡] Adjusted for age, sex, body mass index, smoking status, lingual tonsil volume, and other concurrent multi-level surgery.

Table 4

Multivariate Linear Regression Models: Tonsil Grade & OSA Measures

OSA Measure	Adjusted* Beta Coefficient	95% CI	P-value
Apnea-Hypopnea Index	13.5	(3.5, 23.6)	0.01
Apnea Index	11.3	(3.5, 19.0)	0.01
Oxyhemoglobin Desaturation Index	10.9	(−0.8, 22.6)	0.07
Oxyhemoglobin Nadir	−5.2	(−9.0, −1.3)	0.01
Epworth Sleepiness Scale	1.1	(−1.9, 4.2)	0.44

CI = Confidence Interval.

*Adjusted for age, sex, body mass index, smoking status, and lingual tonsil volume.

Table 5

Multivariate Linear Regression Models: Tonsil Volume & OSA Measures

OSA Measure	Adjusted* Beta Coefficient	95% CI	P-value
Apnea-Hypopnea Index	−0.2	(−2.2, 1.9)	0.89
Apnea Index	1.2	(−1.3, 3.8)	0.34
Oxyhemoglobin Desaturation Index	1.5	(−2.4, 5.4)	0.41
Oxyhemoglobin Nadir	−0.2	(−1.4, 1.0)	0.76
Epworth Sleepiness Scale	0.4	(−0.2, 1.0)	0.17

CI = Confidence Interval.

*
Adjusted for age, sex, body mass index, smoking status, lingual tonsil volume, and other concurrent multi-level surgery.