Short Scientific Communication—Pediatric Otolaryngology

A Pediatric Grading Scale for Lingual Tonsil Hypertrophy

Norman R. Friedman, MD1,2, Jeremy D. Prager, MD1,2, Amanda G. Ruiz1,2, and Eric J. Kezirian, MD, MPH3

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract
Lingual tonsil hypertrophy (LTH) is a common finding for children with residual obstructive sleep apnea (OSA) following an adenotonsillectomy. Secondary to the significant morbidity associated with OSA, identification and treatment of residual OSA are paramount. A dedicated LTH grading scale for children does not exist. The current adult LTH scale is impractical for children. Imaging is not routine for children, since it frequently requires sedation. We present a pediatric LTH grading scale with substantial interrater reliability to facilitate standardization of endoscopy findings and promote outcomes-based research for OSA surgery in children.

Keywords
lingual tonsil hypertrophy, pediatric, children, obstructive sleep apnea, outcomes research, severity or grading

Received May 12, 2015; revised July 20, 2015; accepted July 28, 2015.

Recent research has suggested more limited success of an adenotonsillectomy (T&A) for treating pediatric obstructive sleep apnea (OSA).1 The CHAT study prospectively evaluated OSA outcomes after T&A and reported an overall success of 79%, defined as an apnea-hypopnea index <2 events per hour and an obstructive apnea index <1 event per hour.2 The success of T&A is worse when residual OSA is defined as an apnea-hypopnea index >1 event per hour on postoperative polysomnography rather than parent-reported resolution of symptoms.3 As postoperative polysomnography becomes more common, more children will be identified with residual OSA. Oftentimes, the initial otolaryngic evaluation for residual OSA after T&A includes awake flexible upper airway endoscopy. Increasingly, drug-induced sleep endoscopy is also being utilized.4-9

Grading scales have been established for tonsil size,10 adenoid hypertrophy,11,12 palate position,13 and tongue base obstruction.14 Since lingual tonsil hypertrophy (LTH) is a common finding for children with residual OSA, we have developed a pediatric LTH grading scale (LTS).4,15 Our objective is to introduce the LTS and evaluate its interrater reliability.

Materials and Methods
Grading Scale
Colorado Multiple Institutional Review Board approval was obtained. The LTS was developed with 4 grades assessing the degree of LTH:

Grade 1: None to minimal.
Grade 2: Mild, <50% filling vallecula.
Grade 3: Moderate, >50% effacement of vallecula.
Grade 4: Severe, unable to visualize epiglottis.

The point of observation is standardized by placing the tip of the endoscope at the free edge of the soft palate. The otolaryngologist determines the degree of LTH by assessing to what extent the lingual tonsil occupies the vallecula. Patient cooperation is not uniformly present in the pediatric population, so no special maneuvers are applied. When determining the severity of LTH, one should assign a grade when the tongue is in a neutral position without protrusion. The simplicity of the LTS enables determination even with a brief period of examination under unfavorable conditions.

On the basis of a review of numerous photographs obtained during office and intraoperative flexible transnasal laryngoscopy on pediatric patients, we reached a consensus on the grading scale (Figure 1). Sixteen representative images were selected (4 images for each grade), and an Internet survey was developed with the LTS appearing below each image.

Initial pilot testing with 4 pediatric otolaryngologists evaluated the clarity of instructions and comprehension of

1Department of Otolaryngology, University of Colorado School of Medicine, Aurora, Colorado, USA
2Division of Pediatric Otolaryngology, Children’s Hospital Colorado, Aurora, Colorado, USA
3Department of Otolaryngology–Head and Neck Surgery, Keck School of Medicine of USC, Los Angeles, California, USA

Corresponding Author:
Norman R. Friedman, MD, Department of Pediatric Otolaryngology, Children’s Hospital Colorado, 13123 E 16th Avenue, B455, Aurora, CO 80045, USA.
Email: Norman.Friedman@childrenscolorado.org

DOI: 10.1177/0194599815601403
http://otojournal.org
Five pediatric otolaryngologists then rated all 16 images after receiving electronic written instructions. Data were entered directly into a REDCap database.

Statistical Analysis

Weighted Fleiss kappa statistics were used to evaluate interobserver agreement among raters. The images were assumed to be unordered.

The kappa statistic was interpreted according to guidelines set forth by Landis and Koch. Kappa scores of zero indicate that agreement among observers is no better than if based solely on chance. Kappa scores of 0.01 to 0.19 represent slight agreement; 0.20 to 0.29, fair agreement; 0.40 to 0.59, moderate agreement; 0.60 to 0.79, substantial agreement; and 0.8 to 0.99, near perfect agreement. Any kappa score with an associated P value <.05 was considered statistically significant.

Results

The 16 images were presented to raters in a survey in random order. Overall Fleiss kappa statistic was 0.73 (SD = 0.046, P < .0001), denoting substantial agreement (Table 1). Nine images showed perfect interrater agreement, and 7 showed a maximum grade difference by raters of 1 grade. There were no missing data. Figure 2 provides examples of an image that was reviewed with full concordance among all 5 reviewers and another that had the most nonconcordant responses.

Discussion

A standardized LTH scale for children does not exist. Two standardized LTH scales exist for adults. One was developed for an investigation of pharyngeal reflux, but no interrater reliability was performed. The more recent adult LTH scale, based on observation and imaging, had substantial interrater reliability. Although imaging allows objective measurement of LTH, it is not practical for children.

We proposed a simple and practical tool to document LTH in children. Interrer reliability was substantial. We used the Brodsky palatine tonsil scale as the template, identifying 4 classifications (Figure 1). Based on Yellon et al’s finding that body position has no effect on base of tongue assessment, the LTS may be usable during patient examination in the upright and supine position. A limitation of a purely visual scale is that one could mistakenly identify superficial lingual tonsil hyperplasia as grade 4 if there is a markedly enlarged tongue base. Retroflexion or posterior displacement of the epiglottis can also widen the vallecula and even contribute to upper airway obstruction; however, the trained observer should not confuse epiglottis position with LTH. Preoperative office documentation of LTS facilitates surgical planning. If a child has grade 4 LTH, a surgeon may consider lingual tonsillectomy and/or evaluation with drug-induced sleep endoscopy. The sample size is a potential study limitation, but the interrater agreement was statistically significant for all grades. Despite the substantial agreement of the raters, this short scientific communication demonstrates only the scale’s reliability and does not validate the scale, since there is no comparison to a gold standard. The LTS can be used in further research, and one potential area for investigation is evaluation of the depth of LTH during a lingual tonsillectomy.

Conclusion

We advocate the use of this simple, practical, and reliable visual scale to grade lingual tonsil size. We hope that utilization of the LTS will standardize endoscopy findings and
promote outcomes-based research in many areas, including residual OSA surgery in children.

Acknowledgment
We thank Alexandria M. Jensen for helping with the data analysis.

Author Contributions
Norman R. Friedman, study design, interpretation of the data, writing of manuscript; Jeremy D. Prager, study design, critical revision/evaluation and approval of manuscript; Amanda G. Ruiz, data collection, data analysis, interpretation of data, writing and approval of manuscript; Eric J. Kezirian, interpretation of data, writing and critical revision/evaluation approval of manuscript.

Disclosures
Competing interests: Norman R. Friedman is a member of the American Board of Internal Medicine (ABIM) Board of Directors and the ABIM Internal Medicine Exam Committee. To protect the integrity of board certification, the ABIM strictly enforces the confidentiality and its ownership of ABIM examination content, and Dr Friedman has agreed to keep it confidential. No ABIM examination content is shared or otherwise disclosed in this article.

Sponsorships: None.

Funding source: This study was supported in part by National Institutes of Health / National Center for Research Resources (Colorado Clinical and Translational Sciences Institute grant UL1 RR025780). Contents are the authors’ sole responsibility and do not necessarily represent official National Institutes of Health views.

References


