

# Endoscopic Evaluation of Adenoids: Reproducibility Analysis of Current Methods

Murilo Fernando Neuppmann Feres<sup>1</sup> · Juliana Sato Hermann<sup>2</sup> · Ana Carolina Sallum<sup>2</sup>  
Shirley Shizue Nagata Pignatari<sup>2</sup>

<sup>1</sup>Post-graduate Program on Otolaryngology, Head and Neck Surgery, <sup>2</sup>Division of Pediatric Otolaryngology, Department of Otolaryngology, Head and Neck Surgery, Federal University of São Paulo, São Paulo, Brazil

**Objectives.** To investigate intra- and interexaminers' reproducibility of usual adenoid hypertrophy assessment methods, according to nasofiberendoscopic examination.

**Methods.** Forty children of both sexes, ages ranging between 4 and 14 years, presenting with nasal obstruction and oral breathing suspected to be caused by adenoid hypertrophy, were enrolled in this study. Patients were evaluated by nasofiberendoscopy, and records were referred to and evaluated by two experienced otolaryngologists. Examiners analysed the records according to different evaluation methods; i.e., estimated, and measured percentage of choanal occlusion; as well as subjective and objective classificatory systems of adenoid hypertrophy.

**Results.** Data disclosed excellent intraexaminer reproducibility for both estimated and measured choanal occlusion. Interexaminers analysis revealed lower reproducibility rates of estimated in relation to measured choanal occlusion. Measured choanal occlusion also demonstrated less agreement among evaluations made through the right and left sides of the nasal cavity. Alternatively, intra- and interexaminers reliability analysis revealed higher agreement for subjective than objective classificatory system. Besides, subjective method demonstrated higher agreement than the objective classificatory system, when opposite sides were compared.

**Conclusion.** Our results suggest that measured is superior to estimated percentage of choanal occlusion, particularly if employed bilaterally, diminishing the lack of agreement between sides. When adenoid categorization is used instead, the authors recommend subjective rather than objective classificatory system of adenoid hypertrophy.

**Keywords.** Adenoids, Diagnostic techniques, Endoscopy

## INTRODUCTION

Adenoid hypertrophy is known to be associated with several harmful clinical conditions [1-3]. Due to the relevance of this issue, a great deal of interest has been given to diverse methods of examinations and parameters for identification and evaluation of adenoid hypertrophy [2,4-6].

Among various examination methods, nasofiberendoscopy (NFE) has been currently considered the "gold standard" exam for adenoid evaluation [7]. Moreover, NFE is more effective when identifying adenoid hypertrophy [8], and has been indicated as the main diagnostic tool when adenoidectomy is considered [5].

Therefore, several methods of adenoid size assessment by means of NFE have been introduced [9-16], and largely disseminated [5-8,17-21]. However, several of these diagnostic methods [5-9,12-14,18,19] are subjective, or occasionally, poorly described. Even among researchers that employ objective evaluation methods of the adenoid size [10,11,15-17,20,21], several have failed to perform intra- or interexaminers reproducibility tests and reliability analysis [10,15,17,21].

In view of the relevance of the reliability of measurement tools designed for adenoid hypertrophy evaluation [22], the main

• Received February 27, 2012  
Revision March 30, 2012  
Accepted April 25, 2012

• Corresponding author: **Murilo Fernando Neuppmann Feres**  
Post-graduate Program on Otolaryngology, Head and Neck Surgery, Federal University of São Paulo, Rua Afonso Celso, n°. 982, ap°. 12. Vila Mariana, São Paulo 04119-060, Brazil  
Tel: +55-11-2528-6484, Fax: +55-11-5572-6787  
E-mail: mn-feres@uol.com.br

Copyright © 2013 by Korean Society of Otorhinolaryngology-Head and Neck Surgery.

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

objective of this study was to test 4 of the most usual NFE evaluation methods, concerning their intra- and interexaminers reproducibility [9,11,13,16]. Secondly, this study also intended to verify the relationship between readings recorded from the NFE view of the nasal cavity left and right sides, according to the same assessment methods [9,11,13,16].

## MATERIALS AND METHODS

This research was approved by the Ethics Review Board of the institution it was developed (protocol 0181/08).

Forty children of both sexes, ages ranging between 4 and 14 years, were selected from the Institutional Paediatric Otolaryngology Referral Centre. In order to meet inclusion criteria, patients should have or present complaints of nasal obstruction or oral breathing, with suspected diagnosis of adenoid hypertrophy. Children with syndromes or head and neck malformations were excluded. Subjects with acute infection of the respiratory tract, or with history of previous adenoidectomy, were also dismissed. Informed consent was obtained from all the participants.

The selected sample was then submitted to flexible NFE examination. All the exams were performed after topical anesthesia application (lidocaine 2%) at both nostrils. All exams were recorded, and the digital file derived from the video was edited, so the identification of the patient was kept preserved. The edited clips were then handed to two independent, "blind" examiners, both experienced otolaryngologists, and distinct from the ones involved with the NFE recording. Both examiners were consultant physicians, which have been practicing otolaryngologic specialty for, at least, 5 years.

In order to evaluate the clips, both examiners employed four assessment methods [9,11,13,16]. Two of them [13,16] are designed to categorize adenoid hypertrophy on four levels according to objective [16] (objective adenoid classification [Ob-C]), or

subjective criteria [13] (subjective adenoid classification [Sub-C]). The other two assessment methods [9,11] refer to quantitative measurements of nasopharyngeal obstruction, which could be subjectively estimated [9] (estimated choanal occlusion [ECO]), or objectively measured [11] (measured choanal occlusion [MCO]).

Examiners were oriented to choose the frame sequence that would provide the best view of the adenoid in relation to the choana, obtained from the most distal portion of the inferior turbinate. At these frames, the patient should be performing inspiration exclusively through the nose, with no evidence of the soft palate elevation. The assessment methods (Ob-C, Sub-C, ECO, and MCO) were applied on different periods of time, which permitted truly independent evaluations.

### MCO (%)

In order to employ this method [11], the examiner selected a single clip frame. The selected frame was then converted into a digital file (JPEG format), and MCO was finally calculated by ImageJ [23], an image processing software, as the percentage of the choanal area occupied by the adenoid tissue (Fig. 1).

### ECO (%)

According to this method [9], examiners estimated the degree of nasopharyngeal obstruction relying exclusively upon subjective perception.

### Ob-C

According to this method [16], adenoid hypertrophy is classified according to its anatomical relationship with adjacent structures such as vomer, soft palate and torus tubaris: 1) grade 1, none of the above-cited structures contact with the adenoid tissue; 2) grade 2, the adenoid tissue contacts with the torus tubaris; 3) grade 3, the adenoid tissue contacts with torus tubaris and vomer; 4) grade 4, the adenoid tissue contacts with torus tubaris, vomer and soft palate in resting position.

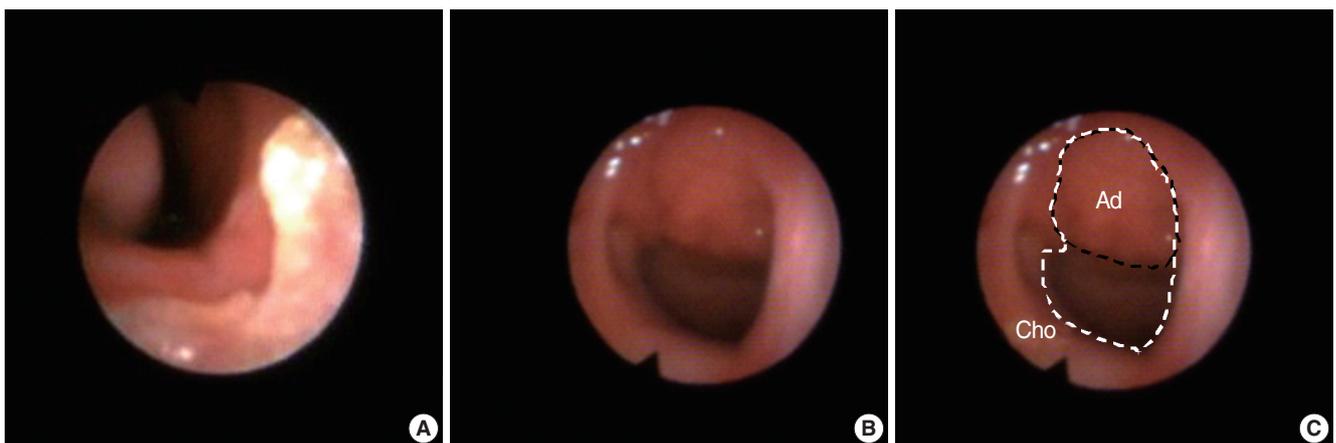


Fig. 1. Final frame selection (B) derived from the clip (A), and posterior calculation of measured choanal occlusion (MCO) (C).  $MCO = (Ad/Cho) \times 100$ .

**Sub-C**

It relies on the examiners' subjective perception, employing the following system of adenoid hypertrophy classification 1) grade 1, adenoid occupying less than 25% of the choanal area; 2) grade 2, adenoid occupying 25-50% of the choanal area; 3) grade 3, adenoid occupying 50-75% of the choanal area; 4) grade 4, adenoid occupying 75-100% of the choanal area [13].

**Statistical analysis**

Reliability of the NFE methods of evaluation was determined by intra- and interexaminers reproducibility analysis. Regarding quantitative variables (MCO and ECO), analysis was accomplished by calculating the intraclass coefficient correlation (ICC), as well as the mean differences between paired readings. Kappa ( $\kappa$ ) coefficient, as well as overall percentage of agreement, which includes agreement occurrences by chance, were employed to analyze reproducibility of the classificatory variables (Ob-C and Sub-C). The relationship between nasal cavity right and left sides' readings was carried out using the same statistical means.

The ICC was interpreted according to Weir [24], which classifies reliability as "poor" ( $ICC \leq 0.20$ ), "reasonable" ( $0.20 < ICC \leq 0.40$ ), "good" ( $0.40 < ICC \leq 0.60$ ), "very good" ( $0.60 < ICC \leq 0.80$ ) or "excellent" ( $0.80 < ICC \leq 1.00$ ). Kappa coefficient was interpreted according criteria described by Landis and Koch [25], whereby the reliability could be characterized as "slight" ( $\kappa \leq 0.20$ ), "fair" ( $0.20 < \kappa \leq 0.40$ ), "moderate" ( $0.40 < \kappa \leq 0.60$ ), "substantial" ( $0.60 < \kappa \leq 0.80$ ) or "almost perfect" ( $0.80 < \kappa \leq 1.00$ ). The level of significance for the statistical tests was 5%.

**RESULTS**

Our research included 20 (50.0%) females and 20 (50.0%) males. Mean age was 9.5 years (range, 4.1 to 14.3 years; standard deviation [SD], 2.4). Clinically, they were all suspected to have adenoid hypertrophy (40/40, 100.00%). Most of the patients complained of mixed (19/40, 47.5%), or exclusively oral breathing (17/40, 42.5%).

According to both evaluations of examiner 1, and examiner 2, MCO mean readings were nearly 70% (71.70% to 73.83%); while ECO mean readings varied from 61.20% to 67.89% (Table 1). Regarding the classificatory parameters (Ob-C and Sub-

C), most of the patients demonstrated grade 3 or 4 adenoid hypertrophy (Table 1). The reproducibility tests were calculated over 71 evaluations (31/40 bilateral records; 9/40 unilateral records), which were randomly ordered before being evaluated by both examiners.

Quantitative diagnostic tools (MCO and ECO), were highly reproducible when employed by the same examiner ( $ICC=0.883, P<0.001$ ;  $ICC=0.885, P<0.001$ , respectively). Interexaminers analysis also showed "excellent" reliability ( $ICC=0.854, P<0.001$ ) for MCO. ECO presented "very good" reliability when performed by distinct examiners ( $ICC=0.728, P<0.001$ ). On average, the same examiner demonstrated 8.09% of variation between paired readings (SD, 8.16%) for ECO; and 4.82% (SD, 5.95%) for MCO. Different examiners demonstrated variation of 10.14% (SD, 10.75%) for ECO, and 5.38% (SD, 6.27%) for MCO.

Regarding the categorical parameters (Ob-C and Sub-C), intraexaminer analysis revealed "substantial" agreement ( $\kappa=0.732, P<0.001$ ) for Sub-C, and "moderate" agreement ( $\kappa=0.457, P<0.001$ ) for Ob-C. Overall percentage of agreement was 81.69% (58/71) for Sub-C and 60.56% (43/71) for Ob-C.

In relation to interexaminers analysis, overall percentage of agreement was 56.33% (40/71) for Sub-C and 50.70% (36/71) for Ob-C. Kappa coefficient calculation showed "fair" agreement for Ob-C ( $\kappa=0.291, P<0.001$ ). Sub-C interexaminers kappa coefficient could not be calculated, since one of the examiners did not classify any patient as grade 1 adenoid hypertrophy. Comparison tests between nasal cavity sides of NFE examination were performed by analyzing exclusively the main examiner (examiner 1) readings. Only patients who had bilateral inspections (31/40) were considered for such analysis.

An "excellent" agreement between sides ( $ICC=0.933, P<0.001$ ) was observed for subjective evaluation ECO. Objective MCO showed "reasonable" agreement ( $ICC=0.404, P=0.010$ ). Moreover, MCO demonstrated larger variation (mean, 10.4%; SD, 11.1%) than ECO (mean, 4.4%; SD, 6.0%) between right and left side readings.

Overall agreement percentage between right and left side evaluations was 90.32% (28/31) for Sub-C, and 67.74% (21/31) for Ob-C. In addition, kappa coefficient revealed "almost perfect" agreement between bilateral evaluations according to Sub-C ( $\kappa=0.842, P<0.001$ ), and "moderate" agreement for Ob-C ( $\kappa=0.550, P<0.001$ ).

**Table 1.** Descriptive analysis of nasofiberoscopy assessment methods, according to examiner 1 (first and second readings), and examiner 2

Examiner	MCO (%)	ECO (%)	Objective adenoid classification				Subjective adenoid classification			
			Grade 1	Grade 2	Grade 3	Grade 4	Grade 1	Grade 2	Grade 3	Grade 4
Examiner 1										
1st reading	73.8±1.8	63.7±2.5	16 (22.5)	11 (15.5)	29 (40.8)	15 (21.1)	5 (7.0)	8 (11.3)	33 (46.5)	25 (35.2)
2nd reading	71.7±2.0	61.2±3.1	13 (18.3)	15 (21.1)	26 (36.6)	17 (23.9)	8 (11.3)	13 (18.3)	26 (36.6)	24 (33.8)
Examiner 2										
	72.1±1.8	67.9±2.2	1 (1.4)	26 (36.6)	41 (57.7)	3 (4.2)	-	11 (15.5)	26 (52.1)	34 (47.9)

Values are presented as mean±SD or number (%).  
MCO, measured choanal occlusion; ECO, estimated choanal occlusion.

## DISCUSSION

The literature reveals large variability concerning NFE methods of adenoid evaluation [5-21]. Among all parameters, four representative diagnostic tools [9,11,13,16] were selected, so their reproducibility could be analyzed. Therefore, further methodological studies are still warranted, so additional assessment methods [10,12,14,15] might as well be evaluated regarding its reproducibility.

The mean age of our sample study is slightly higher (9.5 years) than most of the studies addressing reproducibility of adenoid diagnostic methods [5,6,9,11,16,19,20]. Their sample mean ages varied from 1.25 years [19] to 10.9 years [16]. Any comparison between our results and further literature should consider the differences regarding age groups.

### ECO and MCO

Although both methods showed excellent intraexaminer reliability, interexaminers analysis revealed ECO to have lower rates of reproducibility. In addition, ECO also demonstrated higher intra- and interexaminers differences among paired readings, when compared to MCO. This picture confirms the inherent reliability that is usually expected from objective methods of investigation, and also points to a preferential choice for MCO over ECO, particularly when it comes to the production of scientific evidence. Nevertheless, when MCO is preferred as the method of adenoid evaluation, the authors recommend NFE inspection through both nostrils, since this method revealed lower agreement, and higher variation between opposite sides readings.

In our study, ECO performance was poorer than previously demonstrated [9]. Such study [9] reported maximum (not on average) variation of 10% among examiners. This difference may be related to sampling discrepancies.

Regarding MCO, Demain and Goetz [11] reported only 0.6% of variation between measurements of choanal and nasal areas, whereas in our study the mean variations were 5.38%. Demain and Goetz [11] measurement instrument (nometry over projected transparencies) were distinct from the one we employed (software), what may explain the discrepancy. Yet, both studies revealed acceptable levels of “error” for this method of evaluation (MCO), reinforcing its recommendation over ECO.

### Ob-C and Sub-C

Concerning the objective method (Ob-C), the authors of such a method [16] reported significant degrees of reliability (overall percentage of agreement, 70.48%,  $\kappa=0.71$ ;  $\kappa=0.62$  for medical residents, and  $\kappa=0.83$  for experienced otolaryngologists). A subsequent study [20], confirmed this method to be dependent on the level of experience of the examiner ( $\kappa=0.574$  for medical residents,  $\kappa=0.718$  for experienced otolaryngologists). Overall, our results clearly showed poorer performance. Considering the

differences associated with the level of experience of the examiner [16,20], and the low rates of reliability obtained by our study, the authors recommend specific training strategies, whenever Ob-C is chosen.

Subjective Sub-C presented higher rates of agreement than Ob-C, which is based on objective criteria. Bravo et al. [19] and Ysunza et al. [6] reported even better interexaminers performance (95% of agreement) than the present study. The results provided by this research and available literature [6,19] reinforce the recommendation of this method (Sub-C). Considering Sub-C simplicity and its straightforward use, the authors endorse this method, principally on clinical settings, which demand ease of communication among professionals and prompt diagnosis. In addition, Sub-C method reveals excellent rates of agreement between sides when compared to Ob-C. In that case, the possibility of “one side only” evaluation is recommended if adenoid hypertrophy is the single purpose of NFE examination.

Despite the fact that MCO and Sub-C methods have provided better reliability results, they cannot be accredited as definitive diagnostic methods of adenoid hypertrophy. Diagnostic methods must also include other requirements, such as accuracy, feasibility and, above all, it must positively affect clinical decisions and patient outcome [22].

Future research should then associate reliable (MCO and Sub-C), accurate and practical methods available to a collection of obstructive respiratory symptoms as an effort of systematization of the diagnostic process for adenoid hypertrophy, leading to wise therapeutic management.

Our results suggest that measured is superior to estimated percentage of choanal occlusion, particularly if employed bilaterally, diminishing the lack of agreement between sides. When adenoid categorization is used instead, the authors recommend subjective rather than objective classificatory system of adenoid hypertrophy.

## CONFLICT OF INTEREST

No potential conflict of interests relevant to this article was re-

## ACKNOWLEDGMENTS

This research was financially supported by the State of São Paulo Research Foundation (FAPESP), under the process number 08/53538-0.

## REFERENCES

- Pagella F, Colombo A, Gatti O, Giourgos G, Matti E. Rhinosinusitis

- and otitis media: the link with adenoids. *Int J Immunopathol Pharmacol.* 2010 Jan-Mar;23(1 Suppl):38-40.
2. Farid M, Metwalli N. Computed tomographic evaluation of mouth breathers among paediatric patients. *Dentomaxillofac Radiol.* 2010 Jan;39(1):1-10.
  3. Izu SC, Itamoto CH, Pradella-Hallinan M, Pizarro GU, Tufik S, Pignatari S, et al. Obstructive sleep apnea syndrome (OSAS) in mouth breathing children. *Braz J Otorhinolaryngol.* 2010 Sep-Oct;76(5):552-6.
  4. Jaw TS, Sheu RS, Liu GC, Lin WC. Development of adenoids: a study by measurement with MR images. *Kaohsiung J Med Sci.* 1999 Jan;15(1):12-8.
  5. Mlynarek A, Tewfik MA, Hagr A, Manoukian JJ, Schloss MD, Tewfik TL, et al. Lateral neck radiography versus direct video rhinoscopy in assessing adenoid size. *J Otolaryngol.* 2004 Dec;33(6):360-5.
  6. Ysunza A, Pamplona MC, Ortega JM, Prado H. Video fluoroscopy for evaluating adenoid hypertrophy in children. *Int J Pediatr Otorhinolaryngol.* 2008 Aug;72(8):1159-65.
  7. Kubba H, Bingham BJ. Endoscopy in the assessment of children with nasal obstruction. *J Laryngol Otol.* 2001 May;115(5):380-4.
  8. Bitar MA, Birjawi G, Youssef M, Fuleihan N. How frequent is adenoid obstruction? Impact on the diagnostic approach. *Pediatr Int.* 2009 Aug;51(4):478-83.
  9. Wormald PJ, Prescott CA. Adenoids: comparison of radiological assessment methods with clinical and endoscopic findings. *J Laryngol Otol.* 1992 Apr;106(4):342-4.
  10. Wang D, Clement P, Kaufman L, Derde MP. Fiberoptic examination of the nasal cavity and nasopharynx in children. *Int J Pediatr Otorhinolaryngol.* 1992 Jul;24(1):35-44.
  11. Demain JG, Goetz DW. Pediatric adenoidal hypertrophy and nasal airway obstruction: reduction with aqueous nasal beclomethasone. *Pediatrics.* 1995 Mar;95(3):355-64.
  12. Clemens J, McMurray JS, Willing JP. Electrocautery versus curette adenoidectomy: comparison of postoperative results. *Int J Pediatr Otorhinolaryngol.* 1998 Mar;43(2):115-22.
  13. Cho JH, Lee DH, Lee NS, Won YS, Yoon HR, Suh BD. Size assessment of adenoid and nasopharyngeal airway by acoustic rhinometry in children. *J Laryngol Otol.* 1999 Oct;113(10):899-905.
  14. Monteiro EC, Pilon RR, Dall'Oglio GP. Estudo da hipertrofia adenoidiana: endoscopia X radiografia de nasofaringe. *Rev Bras Otorrinolaringol.* 2000 Jan-Fev;66(1):9-12.
  15. Lourenco EA, Lopes Kde C, Pontes A Jr, Oliveira MH, Umemura A, Vargas AL. Comparison between radiological and nasopharyngolaryngoscopic assessment of adenoid tissue volume in mouth breathing children. *Braz J Otorhinolaryngol.* 2005 Jan-Feb;71(1):23-7.
  16. Parikh SR, Coronel M, Lee JJ, Brown SM. Validation of a new grading system for endoscopic examination of adenoid hypertrophy. *Otolaryngol Head Neck Surg.* 2006 Nov;135(5):684-7.
  17. Wang DY, Bernheim N, Kaufman L, Clement P. Assessment of adenoid size in children by fiberoptic examination. *Clin Otolaryngol Allied Sci.* 1997 Apr;22(2):172-7.
  18. Cassano P, Gelardi M, Cassano M, Fiorella ML, Fiorella R. Adenoid tissue rhinopharyngeal obstruction grading based on fiberendoscopic findings: a novel approach to therapeutic management. *Int J Pediatr Otorhinolaryngol.* 2003 Dec;67(12):1303-9.
  19. Bravo G, Ysunza A, Arrieta J, Pamplona MC. Videonasopharyngoscopy is useful for identifying children with Pierre Robin sequence and severe obstructive sleep apnea. *Int J Pediatr Otorhinolaryngol.* 2005 Jan;69(1):27-33.
  20. Castillo T C, Corssen J C, Breinbauer K H, Namoncura P C. Adenoids assessment using nasopharyngolaryngoscopy: a method validation. *Rev Otorrinolaringol Cir Cabeza Cuello.* 2008;68(2):143-8.
  21. Caylakli F, Hizal E, Yilmaz I, Yilmazer C. Correlation between adenoid-nasopharynx ratio and endoscopic examination of adenoid hypertrophy: a blind, prospective clinical study. *Int J Pediatr Otorhinolaryngol.* 2009 Nov;73(11):1532-5.
  22. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Delineando a pesquisa clinica: uma abordagem epidemiologica. *Porto Alegre: Artmed;* 2008.
  23. Rasband WS. ImageJ [Internet], Bethesda: US National Institutes of Health; c1997-2012 [cited 2013 Jan 15]. Available from: <http://imagej.nih.gov/ij/>.
  24. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005 Feb;19(1):231-40.
  25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977 Mar;33(1):159-74.