Otolaryngology–Head and Neck Surgery (2006) 134, 979-984

ORIGINAL RESEARCH

The effectiveness of tonsillectomy and adenoidectomy in the treatment of pediatric obstructive sleep apnea/hypopnea syndrome: A meta-analysis

Scott E. Brietzke, MD, MPH, and **Daniel Gallagher, MD,** Washington, DC

OBJECTIVE: Present and evaluate the currently available literature reporting on the effectiveness of adenotonsillectomy (T/A) in treating obstructive sleep apnea/hypopnea syndrome (OSAHS) in uncomplicated pediatric patients.

STUDY DESIGN AND SETTING: Systematic review of the literature and meta-analysis of the reduction of the polysomnogram (PSG)-measured Apnea Hypopnea Index (AHI events/hour) resulting from T/A and the overall success rate of T/A in normalizing PSG measurements (%).

RESULTS: Fourteen studies met the inclusion criteria. Mean sample size was 28. All were case series (level 4 evidence). The summary change in AHI was a reduction of 13.92 events per hour (random effects model 95% CI 10.05-17.79, $P < 0.001$) from T/A. The summary success rate of T/A in normalizing PSG was 82.9% (random effects model 95% CI 76.2%-89.5%, $P < 0.001$).

CONCLUSION/SIGNIFICANCE: T/A is effective in the treatment of OSAHS. However, success rates are far below 100%, which could have far-reaching pediatric public health consequences.

EBM rating: B-2a

Blockage of the upper airway during sleep (obstructive sleep apnea/hypopnea syndrome [OSAHS]) has equaled and likely exceeded recurrent tonsillitis as the leading indication for pediatric adenotonsillectomy in the United States.[1] OSAHS is estimated to affect 1% to 2% of all children.[2] Untreated OSAHS is associated with neurobehavioral and cardiovascular dysfunction.[3] The primary cause of pediatric OSAHS is enlargement of the tonsils and adenoids, which then obstruct the airway during the generalized loss of muscle tone that occurs during sleep. OSAHS incidence peaks during early childhood (3-8 years old), mirroring the typical period of tonsil and adenoid enlargement. Removal of the tonsils and adenoids is therefore considered to be the first-line therapy for pediatric OSAHS.

Adenotonsillectomy (T/A) is widely considered to be highly effective in treating OSAHS. The actual effectiveness is not definitively known. Approximately 250,000 pediatric adenotonsillectomies are performed in the United States alone each year with likely the majority for the indication of OSAHS.[4] If even a small percentage (5%) of the procedures are not successful, the potential pediatric public health impact would be enormous.

Polysomnography (PSG) is accepted as the gold standard for evaluation of pediatric OSAHS.[5] However, the associated delay in treatment, high cost, and limited access in some areas have led to infrequent use of this study preoperatively. Likewise, when PSG is used, it is typically for children in high-risk groups (eg, morbid obesity, craniofacial syndromes) or those with persistent symptoms after T/A. These factors have led to the paucity of pre- and postadenotonsillectomy PSG data reported in the medical

From the Department of Otolaryngology, Walter Reed Army Medical Center.

This work supported in part by a CORE Grant from the American Society of Pediatric Otolaryngology.

The opinions and assertions of the authors contained herein are the private views of the authors and are not to be construed as reflecting the views of the Department of Defense or the Department of the Army.

Presented at the Annual Meeting of the American Academy of Otolaryngology–Head and Neck Surgery, Los Angeles, CA, September 25-28, 2005.

Reprint requests: Scott E. Brietzke, MD, MPH, Dept. of Otolaryngology, Walter Reed Army Medical Center, Washington, DC 20307.

E-mail address: sebrietzke@msn.com.

doi:10.1016/j.otohns.2006.02.033

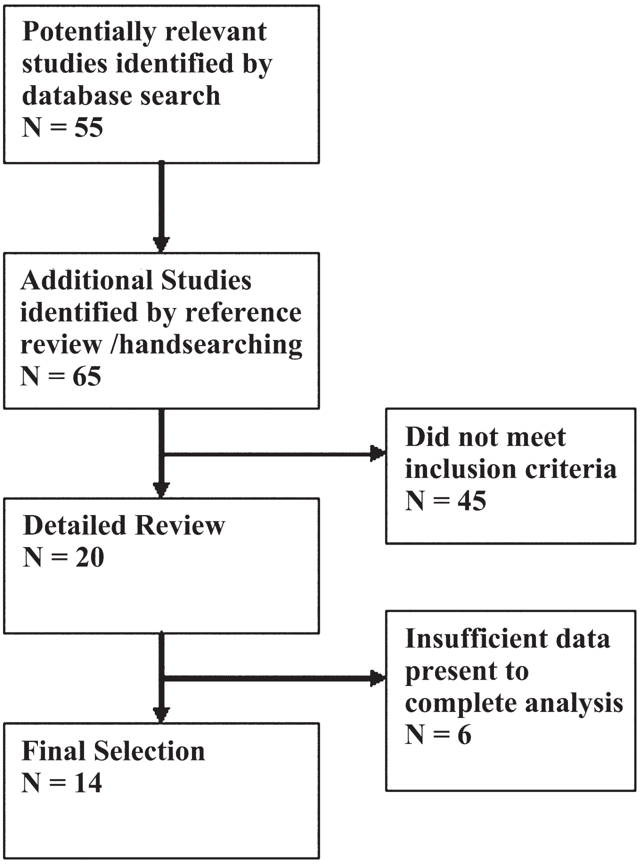

Figure 1 Flow diagram of literature search.

literature. The purpose of this investigation is 1) to systematically review the biomedical literature to identify studies that objectively report on the effectiveness of T/A in treating OSAHS by using PSG data before and after surgery, and 2) to critically evaluate these studies using meta-analysis techniques to begin to objectively assess the effectiveness of T/A in managing pediatric OSAHS.

METHODS

Literature search

A comprehensive literature search was performed using PUBMED, the Cochrane database, and EMBASE as of December 1, 2005 (Fig 1). Search terms used included medical subject headings (MESH) and keywords "tonsillectomy," "adenoidectomy," "outcomes," "indications," "polysomnogram," "polysomnography," "sleep apnea," "obstructive sleep apnea," "apnea," and "hypopnea." Article titles, abstracts, and reference sections were reviewed to determine article eligibility in comparison to inclusion criteria and to identify other relevant studies. The inclusion criteria were as follows: 1) Study subjects must be children less than 18 years old (study did not have to present pediatric data only) and be without significant co-morbidity to include craniofacial syndromes, morbid obesity, or neuromuscular disorders. 2) The surgical treatment must consist of removal of adenoids and both tonsils. 3) Subjects must be evaluated before and after surgery with nocturnal polysomnogram to include at a minimum the measurement of apnea and hypopnea using plethysmography and pulse oximetry with calculation of an Apnea/Hypopnea Index (AHI in events/hr). 4) Study must present sufficient data to allow calculation of the parameters of interest (change in AHI and/or number of patients successfully treated). Non-English language was not used as exclusion criteria. All studies meeting the inclusion criteria were included.

Statistical analysis

Because of the difficulty of choosing a single measure of effectiveness for adenotonsillectomy, 3 separate outcome measures were assessed. The first measure was the difference in mean AHI before and after surgery as reported in the article. The second measure to be evaluated was the percentage of subjects who were found to be successfully treated by surgery as per the PSG criteria for OSAHS as stated in each article (ranging from 1 to 5 events per hour). This was modeled as a binomial variable with N trials and Y successes (% success = Y/N). The standard error of the AHI measurement was calculated from confidence intervals or reported *P* values when not explicitly reported. The third measure to be evaluated included Hedge's Adjusted G. This is a unitless parameter that quantifies the effect size for an intervention as a standardized difference in means while accounting for random error and measurement variation (eg, the variation within PSG measurements) with an additional correction for small ($<$10 subjects) sample size.[6] The Hedge's Adjusted G method was based on the change in mean AHI before and after surgery. Fixed-effects and random-effects models were fitted for each measure; however, the random-effects model was selected as more appropriate because of the significant heterogeneity that was observed among studies. Weights were based on the reported standard error of PSG or the inverse of the sample size. Heterogeneity was formally assessed using a χ^2 test statistic. A priori planned subgroup analyses included evaluating sex and age if possible using meta-regression. Publication bias was assessed using a funnel plot and visual inspection. All statistical analyses were performed with the assistance of computer software (Intercooled STATA V8.2, College Station, TX).

RESULTS

Fourteen articles were selected for final inclusion after detailed review (Fig 1). Each selected article presented pre- and post-surgery AHI data from overnight PSG and/or a percentage of patients successfully treated with T/A based on AHI measurements from overnight polysomnography before and after surgery. The study design was universally a case series (level 4 evidence). Sample size, mean age, and diagnostic criteria for OSAHS were extracted from each study when available (Table

Table 1
Evidence table of studies meeting the inclusion criteria (k = 14)

Authors	Year	Study design	Evidence level	n	Mean age	PSG criteria for OSAHS	Pre-AHI	Post-AHI	Success of T/A (%)
Eliaschar et al[7]	1980	Case series	4	2	6.5	1	6.375	0.125	100
Frank et al[8]	1983	Case series	4	7	NR#	NR	26.2	1.85	NR
Suen et al[9]	1995	Case series	4	26	4.53	5	18.1	4.5	84.6
Nashimura et al[10]	1996	Case series	4	35	4.91	5	13.4	3.37	82.8
Wiet et al[11]	1997	Case series	4	13†	7.5	NR	23	6	NR
Shintani et al[12]	1998	Case series	4	114	4.4	0.5	18.5	7.5	75.4
Wang et al[13]	1998	Case series	4	11	6.7	5	31.5	1.9	100
Bar et al[14]	1999	Case series	4	10	6	1	7.8	1	NR
Nuyens et al[15]	1999	Case series	4	6	6.2	5	9.75	0.5	100
Nieminen et al[16]	2000	Case series	4	21	5.6	1	6.9	0.3	90.5
Avelino et al[17]	2002	Case series	4	17	NR§	1	9.45	0.74	76.4*
Tal et al[18]	2003	Case series	4	36	6.7	5	7.3	1.23	88.9
Mora et al[19]	2003	Case series	4	40	NR‡	NR	26.9	2.6	92.5
Guilleminault et al[20]	2004	Case series	4	17	6.5	1	26	2.3	52.9

NR - not reported.
*Success rate based on tabulated data in manuscript.
†Subset of study population that met inclusion criteria.
‡Range 2-14 years.
#Mean not reported but range-2-13 years.
§Range 2-1 years.

1). The mean sample size among the selected studies was 28 subjects with a pooled mean age of 4.9 years. The mean time period from T/A to post-op PSG was 98 days (range 35-38 days).

The first effect measure of T/A to be analyzed was the mean change in AHI before and after surgery. This data was available in all 14 studies. The mean preoperative AHI among the studies was 16.8 events per hour (range 6.375-26.9 events/hour) and the mean postoperative AHI was 2.42 events per hour (range 0.3-7.5 events/hour). The mean change in AHI after surgery using the random-effects model, which takes into account potential error from both the pre- and postoperative PSG, was a decrease of 13.92 events per hour (95% CI 10.05-17.79 events/hour, $P < 0.001$) (Fig 2). As expected, there was significant heterogeneity amongst the studies (heterogeneity test statistic $\chi^2 = 205.7$, $P < 0.001$).

The next effect measure to be evaluated was the percentage of pediatric patients with OSAHS that were successfully treated (k = 11 studies) with T/A. This was based on pre- and postoperative PSG data and therefore reflects the ability of T/A to normalize abnormal preoperative PSG measurements. The random-effects model estimate for OSAHS treatment success with T/A was 82.9% (95% CI 76.2%-89.5%, $P < 0.001$) (Fig 3). The reader should note that the criteria for T/A success varied with the stated PSG diagnostic criteria for OSAHS per the individual article (range 0.5-5 events/hour) and this summary estimate represents a "best guess" at the effectiveness of T/A, though importantly it is clearly far below 100%. There was also significant heterogeneity among the studies using this outcome measure, as expected ($\chi^2 = 17.169$, $P = 0.009$).

The final effects measure to be modeled was Hedge's Adjusted G (n = 14 studies). This parameter is a standardized (ie, the change in a value divided by its standard error) parameter reflecting the change in the mean value of a measurement before and after an intervention with a correction factor for small sample size. Because Hedge's Adjusted G is a standardized parameter, values greater than 1.0 (ie, when Hedge's Adjusted G is greater than 1.0 the change in the mean measurement of effect before and after the intervention is greater than its standard error) reflect a significant treatment effect due to the intervention after accounting for standard error and sample size.[6] In this analysis the Hedge's G parameter reflects a standardized measure of the change in AHI before and after surgery adjusted for the standard error of the PSG measurements of AHI and adjusted for the small sample size of many of the studies. The random-effects model estimate of Hedge's Adjusted G was 1.43 (95% CI 1.25-1.60, $P < 0.001$) (Fig 4), indicating a highly effective therapy after accounting for standard error and small sample size. There was again significant heterogeneity among the studies as expected ($\chi^2 = 75.16$, $P = 0.001$).

Publication bias was assessed using standard techniques. Publication bias refers to the possibility that studies with negative results, ie, studies showing T/A to be ineffective in treating pediatric OSAHS, were not published. Graphical analysis does not indicate that significant publication bias is likely to be present. The Begg and Mazumdar Test ($P = 0.171$) and Egge's Test ($P = 0.248$) for publication bias were negative. Meta-regression was performed on a limited basis to investigate potential sources of confounding. Regression analysis did not indicate that mean age, PSG criteria, or preoperative AHI significantly impacted the results of the meta-analysis (data not shown).

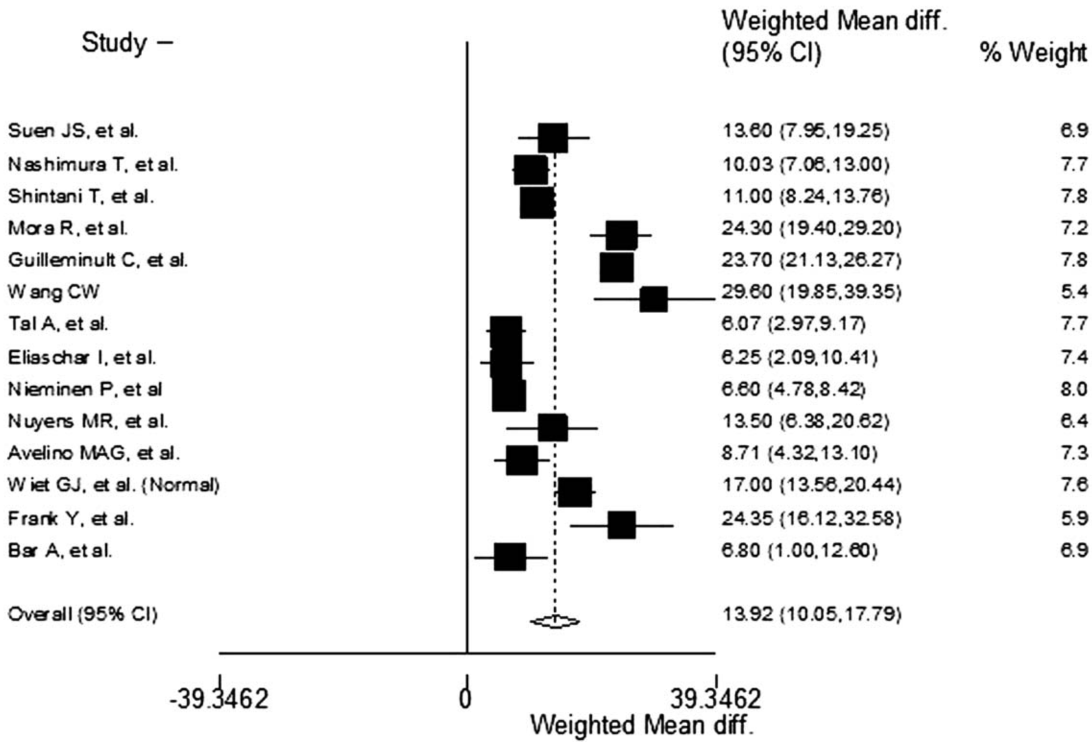

Figure 2 Forest plot for change in AHI after surgery. The size of the box indicates the study's relative weight based on standard error. The dotted line is the overall summary estimate of the change in AHI due to T/A using the random-effects model. The diamond reflects the 95% confidence interval of the summary estimate.

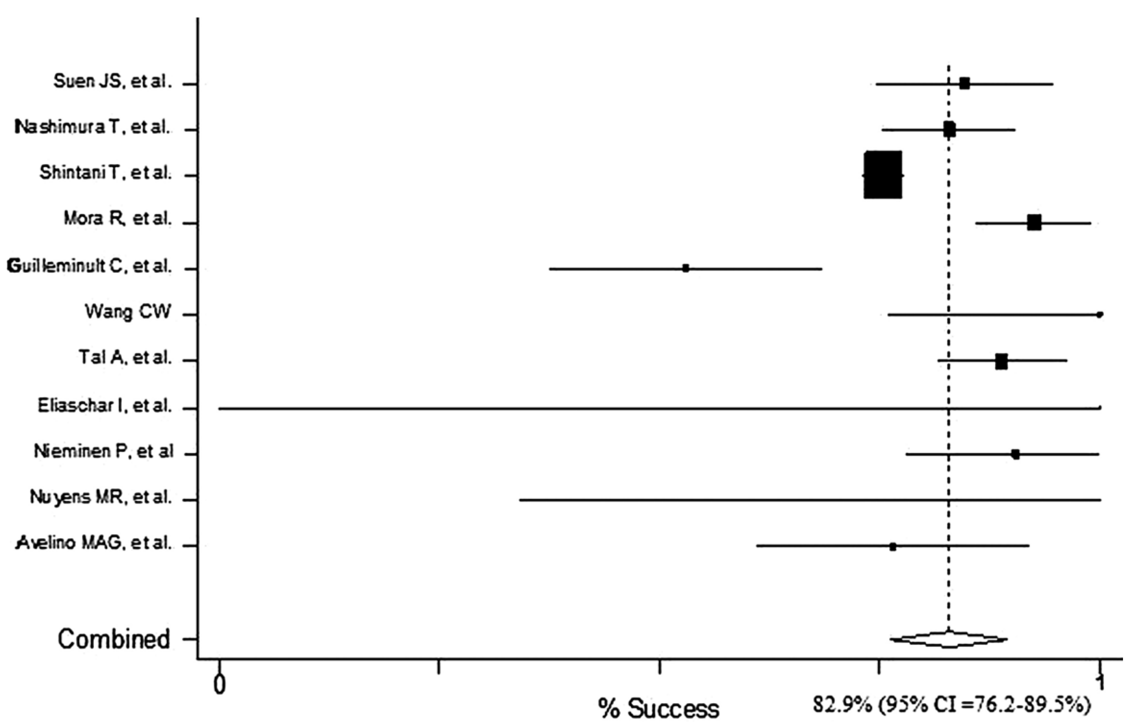

Figure 3 Forest plot for the percentage of patients successfully treated with T/A. The dotted line reflects the overall summary estimate of the percentage successfully treated with T/A using the random effects model. The studies were weighted using the inverse of the sample size. The diamond reflects the 95% confidence interval of the summary estimate.

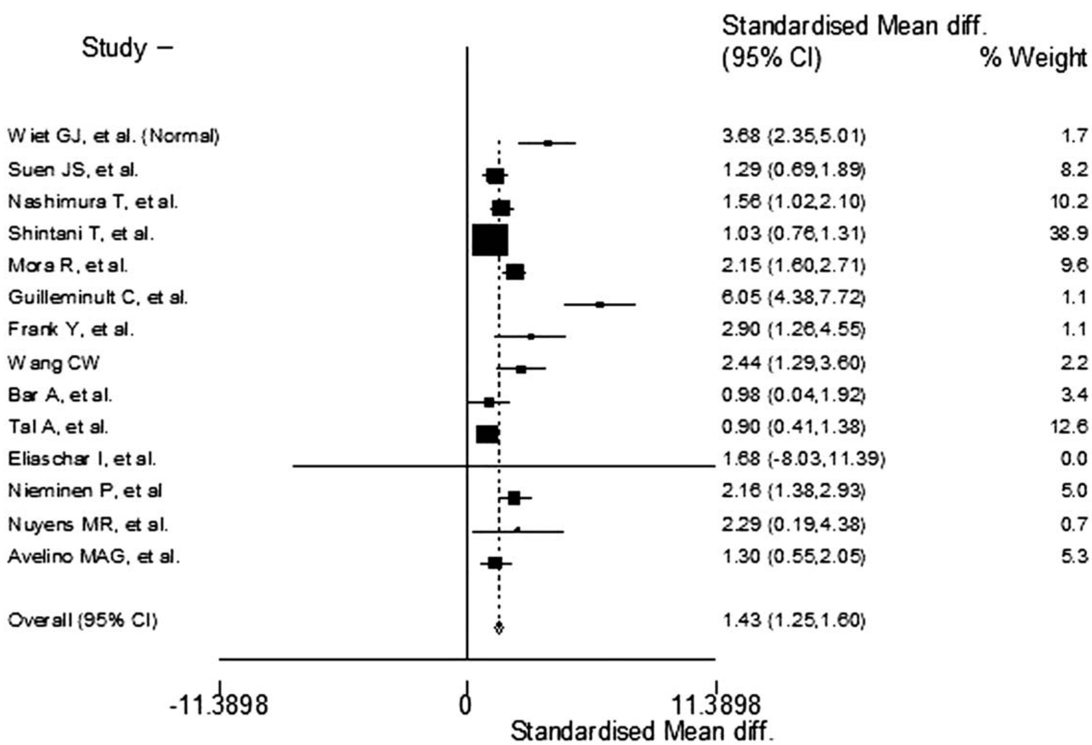

Figure 4 Forest plot for Hedge's Adjusted G Effect measure. Hedge's G is a standardized measure of the effectiveness of an intervention, controlling for standard error and small sample size. Values greater than 1.0 indicate a highly effective intervention (in this case T/A for reducing AHI). The diamond reflects the 95% confidence interval of the summary estimate.

DISCUSSION

This meta-analysis suggests that T/A is effective in treating OSAHS in uncomplicated pediatric patients. The mean AHI and the standardized AHI are significantly decreased after surgery. Additionally, over 80% of patients were successfully treated with T/A using pre- and postoperative AHI measurements to objectively document treatment success. All of these results were highly statistically significant and did not appear to be strongly influenced by possible confounding effects of age, PSG criteria for the diagnosis of OSAHS, or individual sample size.

As expected, the limitation of significant heterogeneity among the studies was demonstrated in the statistical analysis. This arises from several obvious sources. The study populations are likely diverse in ethnicity, body habitus, age range, sex distribution, and other unmeasured factors. Also, the PSG scoring criteria for apneas and particularly hypopneas is controversial and has evolved over time, with significant differences present between earlier and later studies over a 24-year span. Additionally, the threshold for formally diagnosing OSAHS is controversial and has evolved over the years, with earlier studies using a criterion of 5 events per hour and later studies using the more currently accepted threshold of 1 event/hour. In addition, the technology of PSG has also evolved over the years, and it is likely the more recently published studies employed superior PSG equipment with enhanced sensitivity to detect OSAHS than prior studies. Lastly, there are likely differences in the exact surgical technique that was utilized within each of the studies and even from patient to patient.

The limitations of meta-analysis of observational studies have been well documented, and include confounding factors and selection bias.[6] Individual participant data was used when available, and meta-regression analysis was performed in a limited manner to assess for a measurable impact on the results for the some of the above-mentioned factors. Regression coefficients for average age, OSAHS criteria, and preoperative AHI were not found to be significant (data not shown), indicating that there was no evidence of significant confounding of the results from these factors within the limits of meta-regression with only 14 studies.

There appears to be a relative paucity of published studies on the objective effectiveness of pediatric T/A in the treatment of OSAHS despite the high prevalence of pediatric OSAHS and the common performance of T/A for this indication. Evaluation of publication bias did not indicate a high probability of potentially confounding unpublished literature (data not shown).

This study infers that pediatric T/A is far less than 100% effective in treating OSAHS. If true, this fact would have far-reaching public health implications for otolaryngologists. Considering that over 250,000 pediatric T/As are performed in the United States alone with likely the majority being performed for OSAHS,[4] it is reasonable to conclude that residual OSAHS will be present in a large number of children after T/A. Further complicating this scenario are the limitations of the routine clinical evaluation, which has demonstrated poor accuracy in diagnosing pediatric OSAHS and in particular has especially poor accuracy in excluding OSAHS.[21] Therefore, to maintain optimal diagnostic accuracy, clinicians must have a low threshold for the use of objective "gold-standard" polysomnographic testing or suitable screening tests as they are developed. In situations where the probability of residual OSAHS is known to be enhanced, such as obesity, neuromuscular disorders, Trisomy 21, and other craniofacial syndromes, routine postoperative objective testing should likely be employed.[11] In "uncomplicated" pediatric patients routine use of PSG postoperatively is not practical from an access or financial standpoint, yet may be of crucial importance in diagnosing children with residual OSAHS after T/A that would otherwise be considered "cured." Future efforts should include collection of prospective data on larger groups of uncomplicated pediatric patents with homogeneous PSG criteria to further refine any possible predictive factors such as preoperative tonsil/adenoid size, pharyngeal/airway dimensions, age, and race that may assist in the judicious use of PSG to screen patients for residual OSAHS. Additionally, inexpensive noninvasive screening tests that could be used in place of PSG could also have enormous potential benefit. Given the magnitude of this problem, these investigations should be a pediatric public health priority.

CONCLUSION

Pediatric T/A is effective in reducing the severity of objectively measured OSAHS. This finding was consistent despite significant heterogeneity in patient populations and in PSG diagnostic criteria and technology. Pediatric T/A is successful in normalizing PSG measurements in the majority of patients. However, given the magnitude of the problem, the presence of residual OSAHS after T/A in a large number of patients is a significant probability that urgently requires more study and intervention.

REFERENCES

1. Weatherly RA, Mai EF, Ruzicka DL, et al. Identification and evaluation of obstructive sleep apnea prior to adenotonsillectomy in children: a survey of practice patterns. Sleep Med 2003;4:297–307.
2. Brunetti L, Rana S, Lospalluti ML, et al. Prevalence of obstructive sleep apnea syndrome in a cohort of 1,207 children of southern Italy. Chest 2001;120:1930–5.
3. Gozal D, O'Brien LM. Snoring and obstructive sleep apnoea in children: why should we treat? Paediatr Respir Rev 2004;5(Suppl A): S371–6.
4. Rosenfeld RM, Green RP. Tonsillectomy and adenoidectomy: changing trends. Ann Otol Rhinol Laryngol 1990;99:187–91.
5. Ray RM, Bower CM. Pediatric obstructive sleep apnea: the year in review. Curr Opin Otolaryngol Head Neck Surg 2005 Dec;13(6): 360–5.
6. Egger M, Smith GD, Altman DG, editors. Systematic reviews in health care: meta-analysis in health care. London: BMJ Publishing Group; 2001. p. 290–2.
7. Eliaschar I, Lavie P, Halperin E, et al. Sleep apneic episodes as indications for adenotonsillectomy. Arch Otolaryngol 1980;106:492–6.
8. Frank Y, Pollak CP, Weitzman ED. Obstructive sleep apnea and its therapy: clinical and polysomnographic manifestations. Pediatrics 1983;71:737–42.
9. Suen JS, Arnold JE, Brooks LJ. Adenotonsillectomy for treatment of obstructive sleep apnea in children. Arch Otolaryngol Head Neck Surg 1995;121:525–30.
10. Nashimura T, Morishima N, Hasegawa S, et al. Effect of surgery on obstructive sleep apnea. Acta Otolaryngol Suppl 1996;523:231–3.
11. Wiet GJ, Bower C, Seibert R, et al. Surgical correction of obstructive sleep apnea in the complicated pediatric patient documented by polysomnography. Int J Pediatr Otorhinolaryngol 1997;41:133–43.
12. Shintani T, Asakura K, Kataura A. The effect of adenotonsillectomy in children with OSA. Int J Pediatr Otorhinolaryngol 1998;44:51–8.
13. Wang RC, Elkins T, Keech D, et al. Accuracy of clinical evaluation in pediatric obstructive sleep apnea. Otolaryngol Head Neck Surg 1998; 118:69–73.
14. Bar A, Tarasiuk A, Segev Y, et al. The effect of adenotonsillectomy on serum insulin-like growth factor-1 and growth in children with obstructive sleep apnea syndrome. J Pediatr 1999;135:76–80.
15. Nuyens MR, Vella S, Bassetti C, et al. Objektiver nutzen der adenotonsillektomie biem kin. Praxis 1999;88:893–9.
16. Nieminen P, Tolonen U, Lopponen H. Snoring and obstructive sleep apnea in children: a 6-month follow-up study. Arch Otolaryngol Head Neck Surg 2000;126:481–6.
17. Avelino MAG, Pereira FC, Carlini D, et al. Avalicao polissongrafico da sindrome de apneia obsrutiva do sono em criancas, ante e apos adenoamigdatomia. Rev Bras Otorrinolaringol 2002;68:308–11.
18. Tal A, Bar A, Leiberman A, et al. Sleep characteristics following adenotonsillectomy in children with obstructive sleep apnea syndrome. Chest 2003;124:948–53.
19. Mora R, Salami A, Passali FM, et al. OSAS in children. Int J Pediatr Otorhinolaryngol 2003;67(Suppl 1):S229–31.
20. Guilleminault C, Li KK, Khramtsov A, et al. Sleep disordered breathing: surgical outcomes in prepubertal children. Laryngoscope 2004; 114:132–7.
21. Brietzke SE, Katz ES, Roberson DW. Can history and physical examination reliably diagnose pediatric obstructive sleep apnea/hypopnea syndrome? A systematic review of the literature. Otolaryngol Head Neck Surg 2004 Dec;131(6):827–32.