

Adenotonsillectomy for upper respiratory infections: evidence based?

B K van Staaïj, E H van den Akker, G J M G van der Heijden, A G Schilder, A W Hoes

Arch Dis Child 2005;90:19–25. doi: 10.1136/adc.2003.047530

See end of article for authors' affiliations

Correspondence to:
Mrs B K van Staaïj,
University Medical Center
Utrecht, Julius Center for
Health Sciences and
Primary Care, location
Stratenum 6.131,
Universiteitsweg 100,
3584 CG Utrecht,
Netherlands;
B.K.vanStaaïj@med.uu.nl

Accepted 4 May 2004

Background: Despite high rates of (adeno)tonsillectomy for upper respiratory infections in western countries, the medical literature offers the physician little support in deciding which child might benefit from the operation.

Methods: A literature search was performed to identify randomised trials and non-randomised controlled studies into the efficacy of tonsillectomy with or without adenoidectomy in children under 18 years. For the outcomes sore throat episodes, sore throat associated school absence, and upper respiratory infections, pooled estimates of the incidence rate ratios and rate differences with 95% confidence intervals were calculated, assuming a Poisson distribution.

Results: Six randomised trials and seven non-randomised controlled studies on the efficacy of adenotonsillectomy in children were evaluated. For sore throat episodes data for 2483 person-years were available. The pooled risk difference was -1.2 episodes per person-year (95% CI -1.3 to -1.1). For sore throat associated school absence 1669 person-years were analysed. The pooled risk difference was -2.8 days per person-year (95% CI -3.9 to -1.6). For upper respiratory infections 1596 person-years were available. The pooled risk difference was -0.5 episodes per person-year (95% CI -0.7 to -0.3).

Conclusions: All available randomised trials and non-randomised controlled studies into the efficacy of (adeno)tonsillectomy had important limitations. The frequency of sore throat episodes and upper respiratory infections reduces with time whether (adeno)tonsillectomy has been performed or not. (Adeno)tonsillectomy gives an additional, but small, reduction of sore throat episodes, days of sore throat associated school absence, and upper respiratory infections compared to watchful waiting

Tonsillectomy with or without adenoidectomy is one of the most commonly performed surgical procedures in children in western countries.¹ The most common indications are recurrent upper respiratory infections (URIs) and obstructive sleep apnoea. For children with obstructive sleep apnoea due to adenotonsillar hypertrophy, adenotonsillectomy has proven effective.^{2–4} Regarding URIs, however, evidence for its efficacy is limited and no nationally or internationally accepted guidelines on the indications for this procedure are available.^{5,6}

In 1998 two reviews were published on trials of the efficacy of (adeno)tonsillectomy ($T \pm Ads$) for recurrent throat infections in children.^{7,8} Both concluded that the available studies were of poor quality. For several reasons, we wondered whether the current evidence on the effectiveness of (adeno)tonsillectomy is as poor. First, a literature search from Medline with search term "tonsillectomy" produced 5771 hits. Although most of these were non-randomised studies, they may offer important additional evidence. Second, both the Cochrane reviewers and Marshall focused on the reduction of sore throat episodes, but for patients and practising physicians, outcomes such as sore throat associated school absence and upper respiratory infections are also important.^{9–12} Third, the recent report of a second trial by the Pittsburgh group may offer additional evidence.¹³

Therefore, we performed a systematic review to provide a quantitative estimate of the effects of (adeno)tonsillectomy on sore throat episodes, upper respiratory tract infections, and sore throat associated school loss, based on available evidence from randomised trials and non-randomised controlled studies.

METHODS

Study retrieval and selection

A computerised literature search was done in the Medline (Index Medicus 01/1966–06/2003), OldMedline (Index Medicus 01/1963–12/1965), and Cochrane databases for articles containing original data on the efficacy of (adeno)tonsillectomy in children (appendix 1).

Reference lists from identified publications were screened to identify pre-1963 studies. Only articles published in English were retrieved.

Eligibility

Studies were included that met the following criteria: (a) randomised trial or non-randomised controlled study investigating the efficacy of (adeno)tonsillectomy; (b) the control group underwent either no surgery or adenoidectomy only; (c) age at inclusion below 18 years; (d) clinically relevant outcome measures were reported—that is, sore throat episodes and/or upper respiratory infections and/or sore throat associated school loss and/or fever episodes; (e) results were published before June 2003.

Selection of articles

Two reviewers (BvS, EHvdA) independently assessed eligibility of studies. Randomised trials and non-randomised controlled studies were summarised separately.

Abbreviations: $T \pm Ads$, (adeno)tonsillectomy; URI, upper respiratory infection

Table 1 Methods assessment; validity criteria for which bias must be considered likely, and validity and data extraction criteria for which incomplete information hampered assessment

First author	Validity score	Bias considered likely for validity criteria concerning:			Insufficient information for validity criteria concerning:			Data extraction score	Insufficient information for data extraction criteria concerning:		
		Number of satisfied validity criteria (maximum = 10)	Population criteria 1-3	Intervention criteria 4-7	Follow up criteria 8-10	Population criteria 1-3	Intervention criteria 4-7		Follow up criteria 8-10	Number of satisfied data extraction criteria (maximum = 8)	Population criterion 1
Paradise II, three way trial	6		4, 7	8, 9				8			
Paradise II, two way trial	5		4, 6, 7	8, 9				8			
Mawson I and II	4		4, 6, 7	8, 9		5		5	1	3, 5	
McKee II	4		4, 7	9	1, 3	5		5		4	7, 8
Paradise I	2	3	4, 6, 7	8, 9	1, 2			7			7
McKee I	1	1, 2	4, 6, 7	8, 9	3	5		6			7, 8

Randomised trials Methods appraisal

The two reviewers performed the quality assessments of the randomised trials independently. The maximum quality score for each study was 10 for internal validity and 8 for external validity (that is, generalisability) (see appendix 2).¹⁴ For each criterion the reviewers assessed the completeness of the information ("yes" 1 point; "no" 0 points, "unclear" ?). Validity scores were used to rank studies.

Data extraction

Information on patient characteristics (P), interventions (I), the contrast between the interventions (C), and outcomes measured (O) were extracted from all included studies.

Clinical outcomes

The main outcome measures were sore throat episodes, days of sore throat associated school absence, and upper respiratory infections. Results as reported by the authors in the articles were used; no attempt was made to retrieve the original data from the authors as four of six trials were performed more than 15 years ago.¹⁵⁻¹⁸ The incidence of upper respiratory infections for the trials by McKee and colleagues¹⁵⁻¹⁶ was calculated by adding up the episodes of cold, cough, influenza illness, other respiratory illness, and otitis media. For the trial by Mawson and colleagues,¹⁷ the incidence of upper respiratory infections was calculated by adding up the episodes of earache, otitis media, and head cold.

Statistical analysis

Effects on the outcome measures were summarised as risk differences and ratios.

For the outcome parameters sore throat episodes, sore throat associated school absence, and upper respiratory infections in randomised trials, pooled estimates of the rate ratio and rate difference with 95% confidence intervals were calculated. Poisson regression was used assuming that the number of observed episodes followed a Poisson distribution. Incidence rates per person-year were calculated to account for differences in duration of follow up between studies. Random effect estimates were reported because the study results were statistically heterogeneous. To test whether the association between (adeno)tonsillectomy and the outcomes was homogeneous in the two types of analyses (per protocol versus intention to treat), the significance of the respective interaction term was tested. When a significant difference was found, it was tested whether the difference resulted in different treatment effects.

Non-randomised controlled studies

Data extraction, clinical outcomes, and statistical analyses were performed in the same way as in the randomised studies. However, no quality assessment scores were derived and pooled estimates were not calculated.

RESULTS

Study selection

Seventy one studies on the efficacy of (adeno)tonsillectomy for upper respiratory infections in children were identified. Fifty two studies were excluded for the following reasons: 19 were uncontrolled studies; seven included only a before-after treatment comparison; 10 included a control group of healthy controls; and 16 were reviews. Multiple publications were excluded from our analyses.¹⁹⁻²⁴ Hence, six randomised trials¹³⁻¹⁵⁻¹⁸ and seven non-randomised controlled studies¹⁸⁻²⁵⁻³⁰ were included.

Validity criteria of the randomised trials

Internal validity of all the randomised trials was rather poor (maximum score 6 out of 10), whereas the external validity (that is, generalisability) was generally better (minimum score 5 out of 8); see table 1.

Characteristics of the randomised trials (table 2)

The inclusion criteria of the trials varied from mild and non-specific to severe and very strict. Marked differences existed between the inclusion periods of the studies and varied from 6 months to as much as 12 years. In five trials the control group received watchful waiting or non-surgical management, and in one study adenolectomy.¹⁶ Mean age at inclusion varied from 6.0 to 8.1 years. In all trials except one,¹⁶ the loss to follow up was considerable: 8% to 39%. The percentage of switchers from the watchful waiting to the (adeno)tonsillectomy group varied, except in one trial,¹⁶ from 12% to 28%. None of the studies supplied a power analysis.

Characteristics of non-randomised controlled studies (table 3)

Five studies were prospective cohort studies and two were retrospective cohort studies. Main limitations, apart from of course the non-randomised allocation to T±Ads, were that the loss to follow up and the number of children who changed from the watchful waiting to the (adeno)tonsillectomy group were not described.

Efficacy

Outcomes of the randomised trials

The outcome sore throat episodes was studied in all trials, sore throat associated days school absence in five trials, and

Table 2 Characteristics and outcomes of randomised trials

First author	Validity score	Inclusion period	Inclusion criteria	Treatment characteristics and group size	Mean age at inclusion (y)	Outcome assessment	Episode at follow up in T±Ads group (y)			Episode at follow up in control group (y)			Risk difference with 95% CI (episodes/year)	Risk ratio with 95% CI	
							1	2	3	1	2	3			
Paradise three way (2002)	6	12 years	Different criteria for different age groups: - 5 or 6 episodes in past year or 4 in past 2 years (age 3-6 years) - 4-6 episodes in past year or 3 in past 2 years (age 7-15 year)	T±Ads (n=59) versus tonsillectomy (n=58) versus no surgery (n=60)	T±Ads: 7.4 Tonsillectomy: 7.4 Control: 7.4	1, 2, and 3 years	Sore throat episodes Sore throat associated days school absence	1.9	1.7	1.3	2.8	2.9	2.3	-1.0 (-1.3 to -0.7)	0.62 (0.54 to 0.72)
Paradise two way (2002)	5	12 years	As Paradise three way 2002	T±Ads (n=73) versus no surgery (n=78)	T±Ads: 7.4 Control: 7.4	1, 2, and 3 years	Sore throat episodes Sore throat associated days school absence	1.9	1.7	1.5	3.6	2.9	2.4	-1.3 (-1.6 to -1.0)	0.57 (0.49 to 0.66)
Mawson (1967)	4	Not given	Children who would be normally placed on the waiting list for T±Ads	T±Ads (n=202) versus no surgery (n=202)	T±Ads: 6.0 Control: 5.8	1 and 2 years	Sore throat episodes URI episodes school absence	0.7	0.6		2.3	1.7		-1.3 (-1.5 to -1.2)	0.33 (0.29 to 0.39)
McKee II (1963)	4	6 months	≥3 throat infections or URI with cervical adenitis in past 12 months	T±Ads (n=100) versus adenoidectomy (n=100)	T±Ads: 6.7 Ads: 6.5	1 year	Sore throat episodes URI episodes Sore throat associated days school absence	0.3			1.5			-1.1 (-1.4 to -0.9)	0.22 (0.15 to 0.32)
Paradise (1984)	2	11 years	≥7 throat infections in past year or ≥5 throat infections in past two years or ≥3 throat infections in past 3 years	T±Ads (n=43) versus no surgery (n=48)	T±Ads: 8.1 Control: 8.1	1, 2, and 3 years	Sore throat episodes Sore throat associated days school absence	1.2	1.6	1.8	3.1	2.7	2.2	-1.2 (-1.7 to -0.8)	0.55 (0.44 to 0.68)
McKee I (1963)	1	11 months	≥3 throat infections or URI with cervical adenitis in past 12 months	T±Ads (n=231) versus no surgery (n=182)	T±Ads: 6.7 Controls: 6.5	1 and 2 years	Sore throat episodes URI episodes Sore throat associated days school absence	0.4	0.3		2.0	1.0		-1.2 (-1.3 to -1.0)	0.23 (0.19 to 0.28)

Table 3 Characteristics and outcomes of non-randomised controlled studies

First author	Design	Inclusion criteria (study and comparison group)	Treatment characteristics and group size	Age at inclusion (y)	Outcome assessment	Outcome	T ± Ads at follow up	Control group at follow up	Risk difference with 95% CI [% improvement or episodes/year]	Risk ratio (and 95% CI)
Kaiser (1926) (as Kaiser 1924)	Prospective follow up	T±Ads: Children who had obvious diseased tonsils and adenoids and who underwent T±Ads Control group: parents refused T±Ads	T±Ads: 1200 Control: 1200	T±Ads and control: 4 to 7	3 years	Presence of complaints: Frequent sore throat* URI: frequent head colds and ear trouble* Frequent fever attacks*	5% 12% 4%	49% 51% 3%	-44% (-47 to -40%) -39% (-42 to -36%) 2% (0.4 to 3%)	0.11 (0.09 to 0.14) 0.24 (0.20 to 0.28) 1.7 (1.1 to 2.7)
Kaiser (1930) (as Kaiser 1931 and 1940)	Prospective follow up	As Kaiser 1926	T±Ads: 2200 Control: 2200	T±Ads and control: 4 to 7	10 years	Presence of complaints: Frequent sore throat* URI: frequent head colds and ear trouble* Frequent fever attacks*	10% 22% 5%	36% 31% 5%	-25% (-28 to -23%) -9% (-11% to -6%) 0.1% (-1% to 1%)	0.29 (0.25 to 0.33) 0.72 (0.66 to 0.8) 1.0 (0.8 to 1.3)
Monroe (1930)	Prospective follow up	Broad inclusion criteria Author states that T±Ads and control group are comparable	TE: 736 Control: 741	4-13	TE group: 3-18 months Controls: 12 months	Frequent colds and sore throat still present as at inclusion*	14%	58%	-44% (-49% to -40%)	0.24 (0.2 to 0.29)
Roydhouse (1970)	Prospective follow up	T±Ads: recurrent attacks of tonsillitis and other respiratory infections Control group: hospital waiting list	T±Ads: 252 Control: 175	T±Ads and control: 6.1	After 1 and 2 years	Sore throat episodes* URI episodes* Sore throat associated days school absence*	0.4/year 1.35/year 0.87/year	2.1/year 1.36/year 3.8/year	-1.7 (-1.8 to -1.5) 0.02 (-0.2 to 0.2) -3.0 (-3.2 to -2.7)	0.21 (0.18 to 0.24) 1.0 (0.89 to 1.14) 0.23 (0.21 to 0.26)
Roos (1978) (as Roos 1979)	Retrospective; data from database	T±Ads and control group: one doctor's diagnosis of tonsillitis, peritonsillar abscess, or hypertrophy in the year before inclusion	T±Ads: 1950 Control: 2089	T±Ads and control: <13	After 1 year	Doctors visit for respiratory diagnosis	0.57	0.76	-0.19†	0.75†
Hentleff (1981)	Retrospective; data from database	As Roos 1978	T±Ads: 2233 Controls: 2670	T±Ads and control: <13	4 years	Total number of respiratory diagnosis in the 4 follow up years (from database)	0.45/year	0.52/year	-0.07†	0.87†
Paradise (1984)	Prospective follow up; allocation according to parental preference	Inclusion criteria as RCT Paradise 1984, but children were assigned according to parental preference	T±Ads: 52 Control: 44	T±Ads: 7.9 Control: 7.6	After 1, 2, and 3 whole years	Sore throat episodes (definition as RCT Paradise 1984) Sore throat associated school absence (definition as RCT Paradise 1984)	1.51/year 5.4/year	2.88/year 6.2/year	-1.4 (-1.8 to -0.9) -1.0 (-1.8 to -0.2)	0.52 (0.42 to 0.64) 0.84 (0.73 to 0.96)

*Outcome not more precisely defined.

†Baseline differences; CI not calculated.

upper respiratory infections in three trials. Outcomes were assessed at 1, 2, and 3 years. In none of the studies data from children with a follow up of less than one year were included in the analyses.

Pooled estimates

For the outcome sore throat episodes, 2483 person-years were analysed (54% in intervention group). The pooled risk difference was -1.2 episodes per year (95% CI -1.3 to -1.1). The heterogeneity between the studies analysed per protocol versus intention to treat was statistically significant ($p < 0.05$), but no significant treatment effect was observed ($p = 0.80$). The pooled risk ratio for sore throat episodes was 0.49 (95% CI 0.30 to 0.79).

For the outcome sore throat associated school absence, 1669 person-years were available (56% in intervention group). The pooled risk difference was -2.8 days per person-year (95% CI -3.9 to -1.6) and the risk ratio for sore throat associated school absence was 0.50 (95% CI 0.26 to 0.97). No significant effect of the different types of analyses (that is, intention to treat or per protocol) was found.

For the outcome upper respiratory infections, 1596 person-years could be analysed (54% in the intervention group). The pooled risk difference was -0.5 episodes per person-year (95% CI -0.7 to -0.3). The risk ratio for upper respiratory infections was 0.97 (95% CI 0.69 to 1.36). Exclusion of the study by McKee,¹⁶ the only study comparing adenotonsillectomy versus adenoidectomy, yielded similar results.

Outcomes of the non-randomised controlled studies

Except for frequent fever attacks (rate ratio ≥ 1.0), rate ratios for all outcomes (episodes of throat infection, sore throat associated days school absence, upper respiratory infections, and doctor's visit for respiratory diagnosis) were ≤ 1.0 , indicating a beneficial effect of (adeno)tonsillectomy.

DISCUSSION

(Adeno)tonsillectomy reduces the incidence of sore throat episodes by 1.2 episodes per year (95% CI 1.1 to 1.3), sore throat associated school absence by 2.8 days per year (95% CI 1.6 to 3.9), and upper respiratory infections by 0.5 episodes per year (95% CI 0.3 to 0.7).

In contrast with the Cochrane reviewers,⁷ who excluded all trials in which children of the surgical group were randomised to adenotonsillectomy instead of tonsillectomy alone, all randomised trials studying the efficacy of (adeno)tonsillectomy were included in the present meta-analysis. In daily practice most children suffering from recurrent throat infections undergo tonsillectomy combined with adenoidectomy and not tonsillectomy alone; in the Netherlands 90% of tonsillectomies in children are combined with adenoidectomy, in the USA this percentage is 84%, in Canada 75%, and in England 32%. Our meta-analysis shows that randomised trials comparing adenotonsillectomy versus watchful waiting or tonsillectomy alone versus watchful waiting provide similar results.

Our meta-analysis was performed to assess the effects of (adeno)tonsillectomy for upper respiratory infections. The effects of (adeno)-tonsillectomy in children with obstructive breathing during sleep were not considered.²⁻⁴

It is important to realise that all trials had serious methodological limitations, which precludes definite conclusions about the effects of (adeno)tonsillectomy on upper respiratory infections. First, the generalisability of the results of the trials can be questioned, since only a very small proportion of children undergoing T \pm Ads was included in the trials. (Adeno)tonsillectomy is one of the most commonly performed surgical procedures in children in western countries; in 1998, for example, 65/10 000 underwent T \pm Ads in England and 50/10 000 in the United States.¹ Yet

the three Pittsburgh trials^{13 18} included only 233 children in the T \pm Ads group and 186 children in the watchful waiting group, with an inclusion period of respectively 11 and 12 years. Second, all studies had significant loss to follow up. This can be associated with either good or poor outcome. However, in four^{13 17 18} of six studies, information about the children who were lost to follow up was provided, and in these studies the rates of throat infection during the preceding follow up period did not differ significantly from the corresponding rates in the respective treatment groups as a whole. Third, three studies were analysed per protocol.^{15 16 18} These per protocol analyses underestimate the treatment effect as in surgical trials only children of the watchful waiting group with severe complaints can change treatment group, whereas children of the surgical group, who may experience serious complaints, cannot change treatment group.^{15 16 18} Fourth, information bias may be considerable since trials on adenotonsillectomy, as most surgical trials,^{31 32} cannot be performed in a true double blind fashion. None of the trials tried to minimise information bias by choosing an objective outcome measure, such as fever measured daily by a validated thermometer automatically storing data.³³ Fifth, none of the trials provided a power analyses. As all trials, but especially the Paradise trials, included relative few patients, their power may be too low, leading to a type II error.

The pooled risk difference for recurrent throat infections was -1.2 episodes per year (95% CI -1.3 to -1.1). However, in three trials the sore throat episode immediately following the operation was not counted.¹⁵⁻¹⁷ Had these been counted, the differences between the groups would have been smaller.

The pooled risk difference for sore throat associated school absence was -2.8 days per year (95% CI -3.9 to -1.6). In none of the trials, however, was sore throat associated school absence immediately following surgery counted. If these days had been included, the rates would probably not have been different. Thus, although (adeno)tonsillectomy reduces the total number of sore throat episodes by a modest 1.2 episodes per year (95% CI 1.1 to 1.3), the reduction in sore throat associated school absence is even more modest. This indicates that the severity of the throat infections in the children of the control group was likely not serious enough to cause substantial school absence.

The pooled risk difference for upper respiratory infections was only -0.5 episodes per year (95% CI -0.7 to -0.3), indicating that (adeno)tonsillectomy has little effect on the incidence of upper respiratory infections. This is important since several recent studies have shown that many ENT surgeons and general practitioners still regard upper respiratory infections as an indication for (adeno)tonsillectomy.⁹⁻¹²

In all studies children of the control group had more sore throat episodes and more upper respiratory infections than the children of the surgical group. In all studies, however, the children of the control group experienced fewer episodes during the follow up period than before study entry (table 2). This natural decrease of the incidence of throat infections is probably attributable to maturation of the immune system with growing age, whether surgery is performed or not. Regression to the mean could also play a role. As a result, surgery induces an additional reduction of sore throat episodes of only 1.2 episodes per year (95% CI 1.1 to 1.3).

It should be emphasised that the results of all trials are indicative of a difference of a strategy involving (adeno)tonsillectomy and a strategy involving initial watchful waiting, knowing that a proportion of the latter will switch. As in many other surgical trials, the number of switchers was high in most trials. It is very likely that these children have had more throat infections than the children who remained in their original allocated group. This may have led to an underestimation of treatment effect.

The non-randomised controlled studies, except one,¹⁸ show the classical shortcomings of non-experimental studies: incomparability of the study groups at baseline, which leads to confounding by indication.³⁴ In these older studies, techniques that can be used to control for these imbalances of known or suspected risk factors such as multivariate adjustment were not used. In their critical article, Selkirk and Mitchell³⁵ already recognised these problems in 1931. Overtime, the quality of the non-randomised controlled studies has improved. For example, in the older studies by Kaiser and Monroe,²⁵⁻²⁷ inclusion criteria and outcomes are ill defined, while in the more recent studies they are better defined.^{18 29 30} Despite these shortcomings the results of the more recent and better controlled studies^{18 28-30} are surprisingly similar to those of the randomised trials and therefore support evidence on the efficacy of (adeno)tonsillectomy from the trials.

This systematic review shows that all trials and controlled studies have important limitations. Throughout all of the studies the frequency of sore throat infections and upper respiratory infections reduces with time, whether (adeno)tonsillectomy has been performed or not, highlighting the importance of *controlled* studies. Available evidence from both the randomised trials and non-randomised controlled shows that (adeno)tonsillectomy gives an additional, but small, reduction of sore throat episodes, sore throat associated school absence, and upper respiratory infections compared to no surgery.

ACKNOWLEDGEMENTS

We thank A R T Donders for advice and assistance with the statistical analyses.

Authors' affiliations

B K van Staaij, G J M G van der Heijden, A W Hoes, Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, Netherlands

E H van den Akker, A G Schilder, Department of Otorhinolaryngology, Wilhelmina Children's Hospital, University Medical Center, Utrecht, Netherlands

Funding: Dutch Health Care Insurance Board (OG-99-060)

APPENDIX 1: MEDLINE SEARCH

((("tonsillectomy"[mh] OR tonsillectomy[all fields] OR adenotonsillectomy[all fields] OR "Adenoidectomy"[mh] OR Adenoidectomy[all fields]) AND ((placebo[all fields] OR "drug therapy"[sh] OR "therapeutic use"[sh:noexp] OR "random*" [all fields] OR "randomized controlled trial"[pt] OR "Clinical Trials"[mh] OR "Comparative Study"[mh]) OR ("incidence"[mh] OR "mortality"[mh] OR "follow-up studies"[mh] OR "mortality"[sh] OR prognos*[all fields] OR predict*[all fields] OR course[all fields] OR "Population Surveillance"[mh] OR "Remission, Spontaneous"[mh]) OR ("Review Literature"[MH] OR Meta-Anal* OR "meta-analysis"[pt] OR metaanal*) OR ((quantitativ*[tw] OR systematic*[tw] OR methodologic*[tw]) AND (review*[tw] OR overview*[tw])) OR ((("review"[pt] OR review*[tw]) AND ("medline"[tw] OR "cinahl"[tw] OR "embase"[tw] OR "excerpta"[tw] OR "odds ratio"[tw] OR "pooled"[tw] OR "pooling"[tw]))) NOT (letter[pt] OR editorial[pt] OR comment[pt] OR in vitro[mh] OR ("animal"[mh] NOT ("human"[mh] AND "animal"[mh]))) Field: All Fields, Limits: All Child: 0-18 years.

APPENDIX 2

Criteria for the assessment of internal validity

(V1) Was the treatment allocation performed in an unpredictable sequence?

- (V2) Was the treatment allocation concealed (sealed envelopes, allocation by telephone, etc)?
 (V3) Were the groups similar at baseline regarding prognostic indicators and baseline scores?
 (V4) Was the care provider blinded to the treatment (use of a placebo)?
 (V5) Were co-treatments avoided or standardised?
 (V6) Was the compliance rate (in each group) unlikely to cause bias?
 (V7) Was the patient blinded to the allocated treatment?
 (V8) Was the crossover/dropout rate unlikely to cause bias?
 (V9) Was the outcome assessor blinded to the treatment?
 (V10) Was the timing of the outcome assessment in both groups comparable?

Criteria for the assessment of external validity

- (D1) Were the eligibility criteria specified?
 (D2) Were the compared treatments explicitly described?
 (D3) Was information about the method of assessment of outcome measures presented?
 (D4) Were there a short term (immediately after treatment) and a long term follow up measurement?
 (D5) Were adverse effects described?
 (D6) Was sample size for each group described, after allocation and at outcome measurement?
 (D7) Did the analysis include an intention-to-treat analysis?
 (D8) Were point estimates and measures of variability presented for primary outcome measures?

V1 Random (unpredictable) assignments sequence. Methods of allocation using date of birth, date of admission, hospital numbers, or alternation should not be regarded as appropriate.

V2 Assignment generated by an independent person not responsible for determining the eligibility of the patients. This person has no information about the persons included in the trial and has no influence on the assignment sequence or on the decision about eligibility of the patient.

V3 To receive a "yes", groups must be similar at baseline regarding age, frequency of prior episodes, duration of complaints, and severity of complaints.

V4 (Attempt for) blinding described in paper.

V5 Co-interventions should either be avoided in the design or comparable between the index and control group (use of antibiotics, ventilation tubes, adenoidectomy, and attention from researchers, etc).

V6 Treatment should be provided as randomised. Non-adherence (protocol deviation, crossover) is acceptable if it is <15% for both groups separately and <5 % between groups. Qualitative measurement.

V7 (Attempt for) blinding described in paper.

V8 Quantitative measurement. A yes is scored if "non-adherence" and missing data do not lead to substantial bias.
 V9 The reviewer determines (per outcome parameter) when enough information about the blinding is given in order to score "yes".

V10 Timing of outcome assessment should be identical for all intervention groups and for all important outcome assessments.

D1 The reviewer determines if the eligibility criteria are well described.

D2 Adequate description of both the index and control intervention should be given, so that others could replicate the treatment.

D3 Adequate description of the method of assessment of outcome measurements should be given so that others could replicate the study.

D4 Outcome measurements ≤ 1 year after the intervention and outcome measurements at the end of the intervention period. Follow up time and characteristics of effect measurements should be the same in both groups.

D5 Each event should be described and correctly attributed to the allocated treatment (for example, postoperative bleeding, psychological disturbances).

D6 To be presented for each group at randomisation and for the most important outcome assessments.

D7 All randomised patients are reported/analysed for the most important moments of effect measurement (minus missing values) irrespective of non-compliance and co-interventions.

D8 Both point estimates and measures of variability should be presented (to be scored for each important outcome parameter separately). Point estimates are means, medians, modes, etc. Measures of variability are: standard deviations, 95% confidence interval, etc).

REFERENCES

- 1 van den Akker EH, Hoes AW, Burton MJ, et al. Large international differences in (adeno)tonsillectomy rates. *Clin Otolaryngol* 2004;**29**:161-4.
- 2 Shintani T, Asakura K, Kataura A. The effect of adenotonsillectomy in children with OSA. *Int J Pediatr Otorhinolaryngol* 1998;**44**:51-8.
- 3 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.
- 4 de Serres LM, Derkay C, Astley S, et al. Measuring quality of life in children with obstructive sleep disorders. *Arch Otolaryngol Head Neck Surg* 2000;**126**:1423-9.
- 5 American Academy of Otolaryngology-Head and Neck Surgery-Clinical Indicators for Otolaryngology-Head and Neck Surgery. Available from <http://www.entlink.net/practice/products/indicators/tonsillectomy.html> (accessed 17 June 2004).
- 6 Scottish Intercollegiate Guidelines Network. Available from <http://www.sign.ac.uk/guidelines> (accessed 20 May 2004).
- 7 Burton MJ, Towler B, Glasziou P. Tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis. *Cochrane Database Systematic Reviews* 2000;CD001802.
- 8 Marshall T. A review of tonsillectomy for recurrent throat infection. *Br J Gen Pract* 1998;**48**:1331-5.
- 9 Blair RL, McKerrow WS, Carter NW, et al. The Scottish tonsillectomy audit. Audit Sub-Committee of the Scottish Otolaryngological Society. *J Laryngol Otol* 1996;**110**(suppl 20):1-25.
- 10 Donnelly MJ, Quraishi MS, McShane DP. Indications for paediatric tonsillectomy GP versus consultant perspective. *J Laryngol Otol* 1994;**108**:131-4.
- 11 Capper R, Canter RJ. Is there agreement among general practitioners, paediatricians and otolaryngologists about the management of children with recurrent tonsillitis? *Clin Otolaryngol* 2001;**26**:371-8.
- 12 van den Akker EH, Schilder AG, Kemps YJ, et al. Current indications for (adeno)tonsillectomy in children: a survey in the Netherlands. *Int J Pediatr Otorhinolaryngol* 2003;**67**:603-7.
- 13 Paradise JL, Bluestone CD, Colborn DK, et al. Tonsillectomy and adenotonsillectomy for recurrent throat infection in moderately affected children. *Pediatrics* 2002;**110**:7-15.
- 14 van Tulder MW, Assendelft WJ, Koes BW, et al. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine* 1997;**22**:2323-30.
- 15 Mc Kee WJE. A controlled study of the effects of tonsillectomy and adenoidectomy in children. *Br J Prev Soc Med* 1963;**17**:49-69.
- 16 Mc Kee WJE. The part played by adenoidectomy in the combined operation of tonsillectomy with adenoidectomy. Second part of a controlled study in children. *Br J Prev Soc Med* 1963;**17**:133-40.
- 17 Mawson SR, Adlington P, Evans M. A controlled study evaluation of adenotonsillectomy in children. *J Laryngol Otol* 1967;**81**:777-90.
- 18 Paradise JL, Bluestone CD, Bachman RZ, et al. Efficacy of tonsillectomy for recurrent throat infection in severely affected children. Results of parallel randomized and nonrandomized clinical trials. *N Engl J Med* 1984;**310**:674-83.
- 19 Paradise JL, Bluestone CD, Rogers KD, et al. Comparative efficacy of tonsillectomy for recurrent throat infection in more vs. less severely affected children. *Pediatr Res* 1991;**31**:126A.
- 20 Mawson SR, Adlington P, Evans M. A controlled study evaluation of adenotonsillectomy in children. Part II. *J Laryngol Otol* 1968;**82**:963-79.
- 21 Kaiser AD. Effect of tonsillectomy on the general health of twelve hundred children. *JAMA* 1924;**83**:33-7.
- 22 Kaiser AD. The relation of tonsils and adenoids to infections in children. *Am J Dis Child* 1931;**41**:568-81.
- 23 Kaiser AD. Significance of the tonsils in the development of the child. *JAMA* 1940;**115**:1151-6.
- 24 Roos LL Jr. Alternative designs to study outcomes: the tonsillectomy case. *Med Care* 1979;**17**:1069-87.
- 25 Kaiser AD. Tonsillectomy in children. *JAMA* 1926;**87**:1012-15.
- 26 Kaiser AD. Results of tonsillectomy. *JAMA* 1930;**95**:837-42.
- 27 Monroe JD, Volk VK. Study of the effects of 736 tonsillectomies and of 741 controls. *Am J Public Health* 1930;**49**:5-502.
- 28 Roydhouse N. A controlled study of adenotonsillectomy. *Arch Otolaryngol* 1970;**92**:611-16.
- 29 Roos LL Jr, Roos NP, Henteleff PD. Assessing the impact of tonsillectomies. *Med Care* 1978;**16**:502-18.
- 30 Henteleff PD, Roos LL Jr, Roos NP. Indications for tonsillectomy: a 4-yr follow-up evaluation. *Clin Invest Med* 1981;**4**:123-8.
- 31 van der Graaf Y. Clinical trials: study design and analysis. *Eur J Radiol* 1998;**27**:108-15.
- 32 McLeod RS. Issues in surgical randomized controlled trials. *World J Surg* 1999;**23**:1210-14.
- 33 van Staaij BK, Rovers MM, Schilder AG, et al. Accuracy and feasibility of daily infrared tympanic membrane temperature measurements in the identification of fever in children. *Int J Pediatr Otorhinolaryngol* 2003;**67**:1091-7.
- 34 Grobbee DE, Hoes AW. Confounding and indication for treatment in evaluation of drug treatment for hypertension. *BMJ* 1997;**315**:1151-4.
- 35 Selkirk TK, Mitchell AG. Evaluation of the results of tonsillectomy and adenoidotomy. *Am J Dis Child* 1931:9-41.

