Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Burton, Martin J.; Glasziou, Paul P.; Chong, Lee Y ee; Venekamp, Roderick P.

Link to output in Bond University research repository.

Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Review)

Burton MJ, Glasziou PP, Chong LY, Venekamp RP

WILEY

TABLE OF CONTENTS

Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Martin J Burton[1], Paul P Glasziou[2], Lee Yee Chong[1], Roderick P Venekamp[3]

[1]UK Cochrane Centre, Oxford, UK. [2]Centre for Research in Evidence-Based Practice (CREBP), Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Australia. [3]Department of Otorhinolaryngology & Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands

Contact address: Martin J Burton, UK Cochrane Centre, Summertown Pavilion, 18 - 24 Middle Way, Oxford, OX2 7LG, UK. martin.burton@cochrane.nhs.uk.

A B S T R A C T

Background

Surgical removal of the tonsils, with or without adenoidectomy (adeno-/tonsillectomy), is a common ENT operation, but the indications for surgery are controversial. This is an update of a Cochrane review first published in *The Cochrane Library* in Issue 3, 1999 and previously updated in 2009.

Objectives

To assess the effectiveness of tonsillectomy (with and without adenoidectomy) in children and adults with chronic/recurrent acute tonsillitis in reducing the number and severity of episodes of tonsillitis or sore throat.

Search methods

We searched the Cochrane Ear, Nose and Throat Disorders Group Trials Register; the Cochrane Central Register of Controlled Trials (CENTRAL); PubMed; EMBASE; CINAHL; Web of Science; Cambridge Scientific Abstracts; ISRCTN and additional sources for published and unpublished trials. The date of the most recent search was 30 June 2014.

Selection criteria

Randomised controlled trials comparing tonsillectomy (with or without adenoidectomy) with non-surgical treatment in adults and children with chronic/recurrent acute tonsillitis.

Data collection and analysis

We used the standard methodological procedures expected by The Cochrane Collaboration.

Main results

This review includes seven trials with low to moderate risk of bias: five undertaken in children (987 participants) and two in adults (156 participants). An eighth trial in adults (40 participants) was at high risk of bias and did not provide any data for analysis. Good information about the effectiveness of adeno-/tonsillectomy is only available for the first year following surgery in children and for a shorter period (five to six months) in adults.

We combined data from five trials in children; these trials included children who were 'severely affected' (based on the specific 'Paradise' criteria) and less severely affected. Children who had an adeno-/tonsillectomy had an average of three episodes of sore throats (of *any severity*) in the first postoperative year, compared to 3.6 episodes in the control group; a difference of 0.6 episodes (95% confidence interval (CI) -1 to -0.1; moderate quality evidence). One of the three episodes in the surgical group was the 'predictable' one that occurred in the immediate postoperative period.

When we analysed only episodes of *moderate/severe* sore throat, children who had been more severely affected and had adeno-/tonsillectomy had on average 1.1 episodes of sore throat in the first postoperative year, compared with 1.2 episodes in the control group (low quality evidence). This is not a significant difference but one episode in the surgical group was that occurring immediately after surgery.

Less severely affected children had more episodes of moderate/severe sore throat after surgery (1.2 episodes) than in the control group (0.4 episodes: difference 0.8, 95% CI 0.7 to 0.9), but again one episode was the predictable postoperative episode (moderate quality evidence).

Data on the number of sore throat days is only available for moderately affected children and is consistent with the data on episodes. In the first year after surgery children undergoing surgery had an average of 18 days of sore throat (of which some - between five and seven on average - will be in the immediate postoperative period), compared with 23 days in the control group (difference 5.1 days, 95% CI 2.2 to 8.1; moderate quality evidence).

When we pooled the data from two studies in adults (156 participants), there were 3.6 fewer episodes (95% CI 7.9 fewer to 0.70 more; low quality evidence) in the group receiving surgery within six months post-surgery. However, statistical heterogeneity was significant. The pooled mean difference for number of days with sore throat in a follow-up period of about six months was 10.6 days fewer in favour of the group receiving surgery (95% CI 5.8 fewer to 15.8 fewer; low quality evidence). However, there was also significant statistical heterogeneity in this analysis and the number of days with postoperative pain (which appeared to be on average 13 to 17 days in the two trials) was not included. Given the short duration of follow-up and the differences between studies, we considered the evidence for adults to be of low quality.

Two studies in children reported that there was "no statistically significant difference" in quality of life outcomes, but the data could not be pooled. One study reported no difference in analgesics consumption. We found no evidence for prescription of antibiotics.

Limited data are available from the included studies to quantify the important risks of primary and secondary haemorrhage.

Authors' conclusions

Adeno-/tonsillectomy leads to a reduction in the number of episodes of sore throat and days with sore throat in children in the first year after surgery compared to (initial) non-surgical treatment. Children who were more severely affected were more likely to benefit as they had a small reduction in moderate/severe sore throat episodes. The size of the effect is very modest, but there may be a benefit to knowing the precise timing of one episode of pain lasting several days - it occurs immediately after surgery as a direct consequence of the procedure. It is clear that some children get better without any surgery, and that whilst removing the tonsils will always prevent 'tonsillitis', the impact of the procedure on 'sore throats' due to pharyngitis is much less predictable.

Insufficient information is available on the effectiveness of adeno-/tonsillectomy versus non-surgical treatment in adults to draw a firm conclusion.

The impact of surgery, as demonstrated in the included studies, is modest. Many participants in the non-surgical group improve spontaneously (although some people randomised to this group do in fact undergo surgery). The potential 'benefit' of surgery must be weighed against the risks of the procedure as adeno-/tonsillectomy is associated with a small but significant degree of morbidity in the form of primary and secondary haemorrhage and, even with good analgesia, is particularly uncomfortable for adults.

PLAIN LANGUAGE SUMMARY

Surgical removal of the tonsils (tonsillectomy) for chronic or recurrent acute tonsillitis

Review question

This review compared the clinical effectiveness and safety of surgery (removal of the tonsils - tonsillectomy, or adenotonsillectomy - removal of the tonsils and adenoid tissues) against non-surgical management in adults and children with frequent or chronic tonsillitis.

Background

Surgical removal of the tonsils is a commonly performed operation in patients with chronic or recurrent infections of the tonsils (tonsillitis) or the other tissues at the back of the throat (pharyngitis). Sometimes, the adenoid tissues are also removed during the surgery. However, opinions vary greatly about whether or not the benefits of these operations outweigh the risks.

Study characteristics

This review included evidence available up to 30 June 2014. Seven trials with low to moderate risk of bias were included: five in children (987 participants) and two in adults (156 participants). An eighth trial in 40 adults was at high risk of bias and did not provide any data for analysis.

Although some studies in children followed participants for two or three years, reliable information is only available for up to about one year after surgery due to the high number of participants missing follow-up after the first year. Some studies had recruited children who were more severely affected by tonsillitis than other studies (for example, they had tonsillitis more often and with more severe symptoms). Therefore, we grouped the children into 'severely affected' and 'less severely affected' subgroups.

Two studies in adults had a short duration of follow-up (five to six months following surgery).

Key results

We found that in general children affected by recurrent acute tonsillitis may have a small benefit from adeno-/tonsillectomy: this procedure will avoid 0.6 episodes of any type of sore throat in the first year after surgery compared to non-surgical treatment. The children who had surgery had three episodes of sore throat on average compared to 3.6 episodes experienced by the other children. One of the three episodes is the episode of pain caused by surgery.

When it comes to avoiding bad sore throats, children who have more severe or frequent tonsillitis may benefit more from surgery compared to less severely affected children. In less severely affected children the potential benefits of adeno-/tonsillectomy are more uncertain. There are no good quality data for the effects of surgery in the second or later years after surgery.

We did not find enough evidence to draw firm conclusions on the effectiveness of tonsillectomy in adults with chronic/recurrent acute tonsillitis. Evidence is only available for the short term and is of low quality. The data are also difficult to interpret as the studies do not take into account the days of pain that always follow the operation. Based on the two small trials, tonsillectomy seems to result in fewer days of sore throat in the first six months after surgery.

Two of the studies in children said that they could not find a difference in quality of life outcomes and one study could not find a difference in the amount of painkiller that children took to help with their sore throats.

Bleeding immediately after tonsillectomy or in the two weeks following surgery is an important complication. The studies did not provide good information to allow us to assess accurately the risk of these complications.

Quality of the evidence

We judged the quality of the evidence to be moderate for the data on children (this means that further research is likely to have an important impact on how confident we are in the results and may change those results). Quality is affected by a large number of children who are 'lost to follow-up' after the first year of the study. In addition, some children who are assigned to the 'no surgery' group end up having surgery.

The quality of evidence for tonsillectomy in adults in adults is low.

As always, any potential benefits of surgery should be carefully weighed against the possible harms as the procedure is associated with a small but significant degree of morbidity in the form of bleeding (either during or after the surgery). In addition, even with good pain relief medication, the surgery is particularly uncomfortable for adults.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON [*Explanation*]

Tonsillectomy or adenotonsillectomy versus no surgery in children with chronic/recurrent acute tonsillitis

Patient or population: Children with chronic/recurrent acute tonsillitis
Settings: Patients recruited from primary care or specialist centres
Intervention: Tonsillectomy or adenotonsillectomy versus no surgery

Outcomes	Illustrative comparative risks* (95% CI)		No of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	**Corresponding risk**			
	Control	Tonsillectomy or adeno-tonsillectomy			
Episodes of sore throat of any severity Including as 1 episode the period post-surgery) Follow-up: median 1 year	Ranged from **2.9 to 7.7** episodes (median = 3.3, weighted mean = 3.6)[1]	**0.6 fewer episodes** (1.0 to 0.1 lower)	795 (5 studies)	⊕⊕⊕◯ **moderate**[2,3]	The SMD was -0.7 (95% CI -0.9 to -0.5). As a rule of thumb, 0.2 SMD represents small, 0.5 moderate and 0.8 a large effect Data from severely and moderately affected patient subgroups are not shown separately as the results were pooled
Episodes of moderate/severe sore throat (*subgroup of severely affected patients*) Including as 1 episode the period post-surgery) Follow-up: median 1 year	Mean = **1.2** episodes	**0.1 fewer episodes** (0.6 fewer to 0.4 more)	73 (1 study)	⊕⊕◯◯ **low**[3,4,5]	Data from severely and moderately affected patients were not pooled. Tests for subgroup differences were statistically significant

Episodes of moderate/severe sore throat (subgroup of moderately affected patients) Not including as 1 episode the period post-surgery Follow-up: median 1 year	Ranged from **0.2 to 0.9** episodes (median = 0.4 episodes, weighted mean = 0.4 episodes)	**0.8 more episodes** (0.7 to 0.9 more)	491 (3 studies)	⊕⊕⊕◯ **moderate**[2,3]
Sore throat days (including those immediately post-surgery Follow-up: median 1 year	Ranged from **18.9 to 49.1** days (median = 24.0 days, mean = 23.2 days)	**5.1 fewer days** (8.1 to 2.2 fewer)	776 (5 studies)	⊕⊕⊕◯ **moderate**[2,3]
Absence from school or work (standardised to 180 school days or followed up for 5 to 6 months)	Ranged from **5.3 to 6.6 days** (median = 6.3 days, weighted mean = 5.9 days)	**2.3 fewer days** (3.4 to 1.2 fewer)	412 (4 studies)	⊕⊕⊕◯ **moderate**[5]
Complications - results not pooled	Studies reported a number of different complications - not predefined and unclear definitions. See Table 1 for more details			⊕⊕◯◯ **low**[6]
All-cause mortality - not reported	None of the studies reported death as an outcome or complication			⊕⊕◯◯ **low**[6]
Quality of life - results not pooled	Insufficient information to pool results. 2 studies (using different scales) reported no difference			⊕⊕◯◯ **low**[6]

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **SD:** standard deviation; **SMD:** standardised mean difference

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1]Where there were more than three studies, the control group risk shown is the range, median and the weighted mean of the pooled analysis of the control groups of all included studies. The mean shown is the weighted mean of the control groups of all included studies. The assumed risk was calculated from the weighted mean from the meta-analysis.

[2]Borderline decision. All studies were unblinded studies and methods of measurement were not standardised or clearly defined in the studies. One study collected data from patient diaries, but only about 50% of diaries were returned and the data were averages from the returned diaries (Lock 2010). This could be an overestimation of the number of episodes and days with sore throat.

[3]Data were only pooled from studies in children.

[4]Small sample size. Study only reported outcomes for 38 out of 43 and 35 out of 48 patients recruited in the treatment and control group, respectively, and the proportion of patients not included in the analysis is higher in the control group.

[5]There were statistically significant subgroup differences between the moderately affected versus the severely affected groups. Results were not pooled for this outcome.

[6] Results were not clearly reported and could not be analysed. Sample size may be too small to measure reliably mortality and rarer complications.

BACKGROUND

Description of the condition

This is an update of a Cochrane review first published in *The Cochrane Library* in Issue 3, 1999 (Burton 1999) and previously updated in 2009 (Burton 2009).

Sore throat is a common reason for consultation with a doctor in primary care (Ashworth 2004). The symptoms may be caused by infection of the pharynx in general or the tonsils in particular, in which case the terms pharyngitis and tonsillitis may be used respectively. It may be difficult, or impossible, to distinguish accurately between the two conditions, although when the tonsils are red and swollen a diagnosis of tonsillitis is more likely to be made. Doctors disagree on the clinical features that might allow a distinction to be made between tonsillitis and pharyngitis or sore throat.

Although self limiting in the vast majority of patients, the symptoms are inconvenient. As a consequence, antibiotics are frequently prescribed for this condition in general practice (Ashworth 2004). While many patients experience only occasional throat infections with a relatively short duration of illness, a subset of patients suffer from chronic/recurring acute episodes, which causes a considerable disease burden through frequent episodes of sore throat, fever, general illness, sleepless nights, impaired daily functioning and absence from school or work.

Description of the intervention

Surgical removal of the tonsils (tonsillectomy) is one of the most common surgical procedures performed in children (Paradise 1996). Tonsillectomy is performed under general anaesthesia and involves surgical removal of the palatine tonsils from their investing tissues. The risks of surgery include those of the associated general anaesthetic and those specific to the procedure, for example bleeding immediately after surgery or as a result of secondary infection in the 10- to 14-day period after surgery. When the adenoids are also removed, this is called adenotonsillectomy.

How the intervention might work

If a patient has no tonsils, they cannot get tonsillitis. However, they may still suffer from pharyngitis and sore throats. Whilst the palatine tonsils are considered to play an important role in the causation of chronic/recurrent acute throat infections, they are probably not the only factor responsible. As a consequence, tonsillectomy may prevent further throat infections and/or reduce the severity of future throat infections and could therefore considerably improve patients' daily functioning and health-related quality of life. It is unclear whether the removal of adenoids has additional benefits or harms compared to tonsillectomy alone.

Why it is important to do this review

The indications for tonsillectomy (with or without adenoidectomy) are controversial. It is generally accepted that tonsillectomy (with adenoidectomy, if necessary) is indicated in children with obstructive sleep apnoea. The role of the procedure in other types of sleep-disordered breathing is the subject of research and debate (Blackshaw 2014; Venekamp 2014). Although tonsillectomy is also frequently performed in patients who have chronic tonsillitis, recurrent acute tonsillitis or recurrent pharyngitis/'sore throats', opinions vary greatly as to whether or not the benefits outweigh the risks for this indication. We have mentioned the uncertainty surrounding the role of the tonsils (and therefore the effect of their removal) in people with recurrent episodes of tonsillitis/'sore throat'. An additional concern is that a past pattern of behaviour may not continue. Many patients with recurrent symptoms are said to 'grow out' of the problem. A proportion of patients with serious, recurrent symptoms now, will be symptom-free in 12 months' time. It has not been possible to identify specific risk factors that would allow those patients to be identified.

A non-systematic review of tonsillectomy or adenotonsillectomy for recurrent throat infection was published in 1998 (Marshall 1998). This review drew conclusions from trials which either did not fulfil the inclusion criteria for the present systematic review or which appeared to contain significant risk of bias. In the previous version of the current review (Burton 2009), we found four trials undertaken in children (719 participants) and one in adults (70 participants). Good information about the effectiveness of tonsillectomy was only available for (a) children and not adults, and (b) for the effects of surgery in the first year of treatment, not the second year or beyond. At that time we found one ongoing trial on the effectiveness of adeno-/tonsillectomy compared to non-surgical management in children aged 4 to 15 years with recurrent sore throats (the North of England and Scotland Study of Tonsillectomy and Adenotonsillectomy in Children - NESSTAC). In the current update of our review, we included data from this trial (Lock 2010), and one other (Koskenkorva 2013).

OBJECTIVES

To assess the effectiveness of tonsillectomy (with and without adenoidectomy) in children and adults with chronic/recurrent acute tonsillitis in reducing the number and severity of episodes of tonsillitis or sore throat.

As secondary objectives we sought to address the following research questions. In children and adults with chronic/recurrent acute tonsillitis:

- What is the effectiveness of tonsillectomy alone?

- What is the effectiveness of adenotonsillectomy?

- Is there a difference in the effectiveness of tonsillectomy versus adenotonsillectomy?

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs), with no restrictions on publication status or language of publication.

Types of participants

We included children and adults diagnosed by general practitioners or ear, nose and throat (ENT) specialists as having either 'recurrent acute tonsillitis' or 'chronic tonsillitis' and considered them separately.

No microbiological diagnosis was required; we deemed a clinical diagnosis of tonsillitis satisfactory. We considered *recurrent* acute tonsillitis to be more than two distinct episodes in a 12-month period, and *chronic* tonsillitis as symptoms of tonsillitis for a period longer than three months.

Types of interventions

Intervention

Surgical treatment in the form of tonsillectomy, with or without adenoidectomy, by any surgical method (dissection, guillotine, electrocautery, laser, coblation or any other).

Comparator

Non-surgical treatments such as:

1. no treatment (including watchful waiting) with or without analgesics (pain relief) only;
2. repeated short courses of antibiotics; and
3. long-term antibiotics.

We anticipated that some studies may use 'watchful waiting' as the initial management strategy in the non-surgical comparator group, but that some participants allocated to this group may undergo surgery before the end of the study (i.e. cross over into the 'treatment' arm).

Types of outcome measures

We selected three primary outcomes (one of which related to adverse effects of surgery).

Primary outcomes

- Number and severity of episodes of tonsillitis or sore throat
- Number of days with sore throat
- Morbidity and mortality of surgery (measures of morbidity include complications of surgery and number of days with postoperative pain)

Secondary outcomes

- Consumption of antibiotics
- Consumption of analgesics
- Absence or time off work or school
- Quality of life

We sought data on the assessment of outcomes at three months, six months and 12 months, and in the second and subsequent years after randomisation.

In the 2014 update, we added quality of life as an outcome not only because of the availability of data in two included studies, but also to acknowledge the importance of this outcome.

Search methods for identification of studies

We conducted systematic searches for RCTs. We placed no restrictions on language, publication year or publication status. The date of the last search was 30 June 2014, following previous update searches and original searches in 1999.

Electronic searches

We searched the following databases from their inception for published, unpublished and ongoing trials: the Cochrane Ear, Nose and Throat Disorders Group Trials Register; the Cochrane Central Register of Controlled Trials (CENTRAL 2014, Issue 6); PubMed; EMBASE; CINAHL; LILACS; KoreaMed; IndMed; PakMediNet; CAB Abstracts; Web of Science; ISRCTN; ClinicalTrials.gov; ICTRP; Google and Google Scholar. In searches prior to 2013, we also searched BIOSIS Previews 1926 to 2012.

We modelled subject strategies for databases on the search strategy designed for CENTRAL. Where appropriate, we combined subject strategies with adaptations of the highly sensitive search strategy designed by The Cochrane Collaboration for identifying randomised controlled trials and controlled clinical trials (as described in the *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0, Box 6.4.b. (Handbook 2011)). We revised the subject search strategies for the 2014 update of this review. The new search strategies for the major databases are provided in Appendix 1 and have been used in all subsequent update searches. The previous version of the search strategy is provided in Appendix 2.

Searching other resources

We scanned the reference lists of identified publications for additional trials and contacted trial authors where necessary. In addition, we searched PubMed, TRIP database and Google to retrieve existing systematic reviews relevant to this systematic review, so that we could scan their reference lists for additional trials. We searched for conference abstracts using the Cochrane Ear, Nose and Throat Disorders Group Trials Register.

Data collection and analysis

Selection of studies

At least two review authors independently screened titles and abstracts obtained from the database searches at different stages of the original review and subsequent updates. Similarly, at least two of the three review authors independently reviewed the full text of the potentially relevant titles and abstracts against the inclusion and exclusion criteria. We resolved differences by discussion.

Data extraction and management

At least two of the three review authors independently extracted data from the included studies using standardised forms. We resolved disagreements by discussion.
We also noted possible variations in the definition of outcomes between studies, and where possible, standardised these across studies before meta-analysis. See Included studies and Appendix 3 for more details.

Assessment of risk of bias in included studies

At least two review authors independently assessed the methodological quality of the included trials. We resolved any disagreements by discussion. We took the following taken into consideration, as guided by the *Cochrane Handbook for Systematic Reviews of Interventions* (Handbook 2011):

- sequence generation;
- allocation concealment;
- blinding;
- selective outcome reporting;
- incomplete outcome data;
- other sources of bias.

We used the Cochrane 'Risk of bias' tool in Review Manager 5 (RevMan 2014), which involves describing each of these domains as reported in the trial and then assigning a judgement about the adequacy of each entry: 'low', 'high' or 'unclear' risk of bias. Results of the 'Risk of bias' assessment are presented in a 'Risk of bias' summary and a 'Risk of bias' graph.

Measures of treatment effect

We analysed data according to the intention-to-treat principle, whereby all participants are analysed in the groups to which they were randomised. Where data were missing or there was loss to follow-up, we conducted an available case analysis.
We expressed the pooled measures of treatment effect for number and severity of episodes of tonsillitis or sore throat and number of days with sore throat as mean differences with 95% confidence intervals (CIs).

Unit of analysis issues

We identified no studies with non-standard designs, such as crossover and cluster-randomised trials.

Dealing with missing data

We planned to seek key unpublished information missing from reports of included studies, if necessary, by contacting the authors of the included trials. No imputations for missing data were pre-planned, apart from standard calculations to obtain SD values for continuous data as detailed in the *Cochrane Handbook for Systematic Reviews of Interventions* (Handbook 2011).

Assessment of heterogeneity

We assessed clinical heterogeneity (which may be present even in the absence of statistical heterogeneity). We reviewed the included trials for potential differences between studies in terms of the types of participants recruited, interventions or control used and how outcomes were measured and/or reported.
We assessed statistical heterogeneity by visual inspection of the forest plots and statistical tests for heterogeneity. We assessed heterogeneity using the Chi^2 test (with a significance level set at P value < 0.10) and the I^2 statistic, which calculates the percentage of variability which is due to heterogeneity rather than chance, with I^2 values over 50% suggesting substantial heterogeneity (Handbook 2011).
Where there was heterogeneity and a sufficient number of studies was present, we conducted sensitivity analyses based on risk of bias and carried out pre-specified subgroup analyses. Assessments of potential differences in effect sizes between subgroups were based on the Chi^2 tests for heterogeneity statistics between subgroups. We used a fixed-effect meta-analysis where no heterogeneity was present. If no sensitivity analysis completely resolved statistical heterogeneity then we employed a random-effects (DerSimonian and Laird) model to provide a more conservative estimate of the effect.

Assessment of reporting biases

We planned to compare the outcomes reported in the trial against the protocol for the studies whenever possible to assess for reporting bias. If a sufficient number of studies had been available,

we would have conducted more formal assessments using funnel plots.

Data synthesis

We examined the characteristics of participants, interventions and comparisons and compared them between studies. If variations in the intervention or comparator had been found, we would have investigated these via subgroup analysis or even as different comparisons (i.e. not pooled if these were sufficiently different).

We combined the results of individual studies in a meta-analysis where we judged participants, interventions, comparisons and outcomes to be sufficiently similar, to ensure that an answer would be clinically meaningful. We calculated treatment differences with the Mantel-Haenszel method using a fixed-effect or random-effects model determined by the presence or otherwise of significant heterogeneity.

For dichotomous data, we analysed treatment differences as risk ratio (RR) with 95% confidence interval (CI) calculated by the Mantel-Haenszel method using a fixed-effect model. For continuous data we analysed mean differences (MD) with 95% CIs.

We planned to assume that baseline risk is typically either (a) the median of the risks of the control groups in the included studies, this being used to represent a 'medium-risk population' or, alternatively, (b) the average risk of the control groups in the included studies is used as the 'study population' (Handbook 2011). Should further studies be added in future updates it may also be appropriate to consider assumed baseline risk in (c) a low-risk population, and (d) a high-risk population.

Subgroup analysis and investigation of heterogeneity

We planned to analyse data for adults and children separately. We planned to consider the following subgroup analysis, if sufficient data were available:

- Children who were deemed to be 'severely affected' based on a specific set of criteria often referred to as the 'Paradise criteria' *versus* less severely affected children with recurrent acute tonsillitis.

Sensitivity analysis

When this review was first undertaken, there was no evidence to suggest that variations in surgical treatment (in the form of tonsillectomy with or without adenoidectomy) or in the specific surgical method (dissection, guillotine, electrocautery, laser, coblation or any other) were factors affecting the outcomes.

In one of the studies the outcomes of the tonsillectomy alone and adenotonsillectomy groups were compared (Paradise 1984). Finding no "large or statistically significant" differences between them, the data sets were pooled and reported as a single 'surgical' group. It is, however, not clear whether the sizes of the two original subgroups were large enough to detect any true difference between the

effects of tonsillectomy alone and adenotonsillectomy. Some part of the effect of 'surgery' could be due to removal of the adenoids. Therefore we planned to address these assumptions by sensitivity analysis if data were available. In further sensitivity analysis, we planned to assess the impact of studies with high risk of attrition bias on the outcomes

GRADE and 'Summary of findings'

We used the GRADE approach to rate the overall quality of evidence for each outcome. The quality of evidence reflects the extent to which we are confident that an estimate of effect is correct and we applied this in the interpretation of results. There are four possible ratings: high, moderate, low and very low. A rating of high quality of evidence implies that we are confident in our estimate of effect and that further research is very unlikely to change our confidence in the estimate of effect. A rating of very low quality implies that any estimate of effect obtained is very uncertain.

The GRADE approach rates evidence from RCTs that do not have serious limitations as high quality. However, several factors can lead to the downgrading of the evidence to moderate, low or very low. The degree of downgrading is determined by the seriousness of the these factors:

- study limitations (risk of bias);
- inconsistency;
- Indirectness of evidence;
- imprecision; and
- publication bias.

We included a 'Summary of findings' table (Summary of findings for the main comparison), constructed according to the recommendations described in Chapter 10 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Handbook 2011). We used the GRADE considerations to assess the quality of the body of evidence for each primary outcome, and to draw conclusions about the quality of evidence in the review.

R E S U L T S

Description of studies

See: Characteristics of included studies; Characteristics of excluded studies; Characteristics of ongoing studies.

Results of the search

This is an update of a Cochrane review first published in 1999, Burton 1999, which was previously updated in 2009 (Burton 2009).

With the updated searches (to 30 June 2014) we retrieved a total of 3285 additional records. Removing duplicates left 1568 records.

After screening titles and abstracts, we identified five additional potentially relevant studies (Clayburgh 2011; Koskenkorva 2013; Lock 2010; Naiboǧlu 2010; Stalfors 2012). We excluded three as they were non-randomised cohort studies (Clayburgh 2011; Naiboǧlu 2010; Stalfors 2012), leaving two additional trials for consideration (Koskenkorva 2013; Lock 2010). In addition, we decided to include a study that had not previously been included (Stafford 1986); although this study had not reported any data that could be used for meta-analysis and the randomisation method was unclear, it did meet the inclusion criteria from our protocol. Searches in April 2008 retrieved 2746 references, of which we considered nine studies to be possibly relevant. We excluded one as it was a non-randomised cohort study (Kaiser 1930), leaving eight trials for consideration (Alho 2007; Mawson 1967; McKee 1963; Paradise 1984; Paradise 2002a; Paradise 2002b; Roydhouse 1970; van Staaij 2004). The two Paradise 2002 studies are reported in the same paper, but in such a manner that they can be considered as two separate studies (see further details below) (Paradise 2002a;

Paradise 2002b).

Of 10 potentially relevant trials, three were excluded in 1999 because it was not clear whether the included children had suffered from recurrent acute or chronic tonsillitis, as trial entry criteria were poorly defined (Mawson 1967; McKee 1963; Roydhouse 1970). For example, it was possible to be included in the Mawson study by experiencing cervical adenitis (inflamed glands in the neck) alone (Mawson 1967). This left seven RCTs that satisfied all our inclusion criteria (Alho 2007; Koskenkorva 2013; Lock 2010; Paradise 1984; Paradise 2002a; Paradise 2002b; van Staaij 2004). We identified no additional trials after scanning the reference lists of the included studies and relevant systematic reviews. We identified an additional trial, published in 1986, through correspondence with other experts in the field (Stafford 1986), and we have identified one ongoing trial in adults (NATTINA 2014).

Figure 1 shows a flow chart of the number of studies found in the search and included or excluded in the review process.

Figure 1. Process for sifting search results and selecting studies for inclusion

| 5 studies included in the previous version of the review | | 3285 records identified through database searching | 1 additional record identified through other sources |

1569 records after duplicates removed

1569 records screened → 1563 records excluded

6 full-text articles assessed for eligibility → 3 full-text articles excluded, with reasons

3 NEW studies included 2008-2014

8 studies included in qualitative synthesis

7 studies included in quantitative synthesis (meta-analysis)

Included studies

Details of the methods, participants, interventions and outcomes of the included studies are shown in the table of Characteristics of included studies.

Design

We included a total of eight unblinded randomised controlled trials, randomising 987 children and 196 adults, in the review.

Participants and settings

Children

Five of the studies included only children (Lock 2010; Paradise 1984; Paradise 2002a; Paradise 2002b; van Staaij 2004). Four were published after 2000; one study was conducted between 1971 and 1982 (Paradise 1984).

The severity of symptoms required to enter the trials differed significantly. The participants in Paradise 1984 were only eligible if they were severely affected by tonsillitis; their episodes of throat infection met strict, pre-defined standards in each of four categories: (1) frequency of occurrence (seven or more episodes in the preceding year, or five or more in each of the preceding two years, or three or more in each of the preceding three years); (2) clinical features (each episode had to be characterised by specific clinical features); (3) treatment (they had to have been treated with antibiotics when streptococcal infection was proven or suspected); and (4) documentation (each episode had to have been be documented in a clinical record). Meanwhile, the other four studies were less strict in their criteria for severity or frequency of tonsillitis and included moderately affected children (Lock 2010; Paradise 2002a; Paradise 2002b; van Staaij 2004). The van Staaij 2004 trial specifically excluded the type of severely affected child included in the Paradise 1984 study and their participants were likely to be more similar to the children included in the Paradise 2002a; Paradise 2002b trials. In addition, the Lock 2010 trial also included children with less strict criteria regarding the number and reporting of recurrent sore throats before trial entry (four or more episodes of sore throat within each of two years or six or more episodes of sore throat within one year).

All the 'Paradise' studies, which were conducted in the United States, included children aged 3 to 15 (Paradise 1984; Paradise 2002a; Paradise 2002b). Randomisation was stratified into three age groups. Lock 2010 was conducted in the UK and included children aged 4 to 15. van Staaij 2004 was conducted in the Netherlands and included a younger age group (three to eight years old).

All these studies were funded by the relevant national health research or insurance bodies of the respective countries.

Adults

Three studies conducted in adults met the review inclusion criteria (Alho 2007; Koskenkorva 2013; Stafford 1986), but only two of these reported outcomes selected for this review (Alho 2007; Koskenkorva 2013). Both of these were conducted in the Oulu region of Finland and no external source of funding or competing interests were declared.

Alho 2007 recruited participants aged 15 years and above between 2001 and 2005, with proven group A streptococcal pharyngitis. Participants had to have documented recurrent episodes of pharyngitis due to group A *Streptococcus* (three or more episodes in six months or four episodes in 12 months); the symptoms and signs had to be "typical of streptococcal pharyngitis" and severe enough for the patient to seek medical attention. At least one episode had to be microbiologically proven by culture or rapid antigen test.

Koskenkorva 2013 was conducted between 2007 and 2010. This study included patients aged 13 and above who were referred for tonsillectomy because of recurrent pharyngitis (three or more episodes in the previous 12 months).

However, participants from both studies had a similar number of episodes of acute pharyngitis diagnosed by a physician before enrolment (a mean of 3 and 3.5 episodes in the last six months and five episodes in the last 12 months).

Stafford 1986 was conducted in the UK, probably in the early 1980s. The study "randomly" allocated 40 patients referred by their general practitioners. Patients over 16 years old were included if they had a history of at least four episodes of tonsillitis per year in the two preceding years, each of these episodes consisting of sore throat, dysphagia, fever and general malaise for at least three days.

Interventions and comparisons

Children

Participants in the treatment arm of the van Staaij 2004 trial received adenotonsillectomy, while those in Lock 2010 could receive either tonsillectomy (dissection or bipolar diathermy according to surgical preference) or adenotonsillectomy (tonsillectomy with adenoid curettage). The control arms of both studies received "standard" medical treatment.

The Paradise studies had treated tonsillectomy and adenotonsillectomy as different intervention arms. In Paradise 1984, the results were reported together. Paradise 2002a and Paradise 2002b were

two trials that ran in parallel. Paradise 2002b included children with one or more indications for adenoidectomy (obstructing adenoids or history of recurrent or persistent otitis media) assigned randomly into adenotonsillectomy or control non-surgical intervention (two-arm trial). Paradise 2002a only included patients with no indications for adenoidectomy and they were randomly assigned to tonsillectomy, adenotonsillectomy or control (three-armed trial). In the Paradise studies, the protocol for managing suspected sore throat involved taking a culture and prescribing penicillin if the culture was positive.

Adults

The participants in the treatment arm of the two studies in adults received a tonsillectomy as soon as possible after randomisation. All patients had tonsillectomy (total capsular tonsillectomy using blunt or diathermy dissection), except one (who received adenotonsillectomy in Koskenkorva 2013). Both studies used "watchful waiting" as the control group. The control groups 'waited' for between three and six months in Alho 2007 and between five and six months in Koskenkorva 2013.

Patients from both arms were advised to see their general practitioner (GP) whenever they had acute symptoms suggestive of pharyngitis. No specific treatment protocol was stated, although participants were also given information about the study for their GP, along with instructions about examining their ear, nose and throat status, how to take cultures and obtain the relevant blood samples.

In the Stafford 1986 study, participants in the surgery group underwent tonsillectomy within six weeks of trial entry. In the non-surgical group, the participants' general practitioners (GPs) prescribed an eight-day course of antibiotics for each episode of tonsillitis in the following year. Patients in this group were allowed to "opt out" and have surgery at any time. Patients who had surgery were reviewed every six months, up to the 18th month post-surgery. Patients in the antibiotics group were reviewed every three months in the outpatient clinic, and their GPs were sent study protocols and questionnaires, which they completed and returned to the investigators whenever the patients sought treatment (and received antibiotics) for tonsillitis. A final review was carried out between 18 and 24 months after entry to the trial.

Outcomes

The Paradise studies followed up participants over three years. A standardised in-person or telephone inquiry was made bi-weekly concerning the day to day occurrence of a specific list of symptoms and events, and standardised clinical examination was performed every six weeks and at the time of respiratory illness (Paradise 1984).
Both Lock 2010 and van Staaij 2004 collected data through patient diaries. In van Staaij 2004, parents performed additional daily temperature measurements using an electronic thermometer that stored data, and these data were collected along with data from diaries and quality of life questionnaires during follow-up visits at 3, 6, 12, 18 and 24 months. Participants in Lock 2010 were supposed to return the four-weekly diary by post, but telephone prompt interviews were introduced "part way through the study" due to low response rates to the health diary.

Measures of sore throat and sore throat days

Participants in two studies in adults were instructed to see their GP when they had signs and symptoms suggestive of pharyngitis, and instructions were given about collection of data from medical examination and throat culture (Alho 2007; Koskenkorva 2013). Data were also collected by symptom diaries.
All studies reported the number of episodes of sore throat and number of days with sore throat, except for Stafford 1986. However, there were variations in how studies defined what is considered a sore throat episode or a sore throat day, and whether the episode of pain and days of pain following surgery were included in their report of number of episodes of sore throat and sore throat days.
Based on the data reported, for children we were able to standardise the report of sore throat episodes to ensure all studies in children included the episode due to the surgery. No standardisation was required for sore throat days in children as all the included studies already included the postoperative days in their data sets.
The situation was different in adults. Both included studies excluded the episode of postoperative sore throat in those participants undergoing surgery and any days of sore throat due to the surgery.

Definition of sore throat episodes and days

The definition of what is considered an episode of sore throat and sore throat days varied:
- In the Paradise studies, results for episodes of 'throat infection' are reported in four ways: (a) 'moderate or severe' (based on a scoring system), (b) streptococcal, (c) 'counting' (characterised by one or more qualifying clinical features of episodes used in determining trial eligibility), and (d) 'all combined' (Paradise 1984; Paradise 2002a; Paradise 2002b).
- In the van Staaij 2004 trial, an episode 'throat infection' was defined as sore throat or pain or difficulty swallowing *combined with* fever, whilst an episode of 'sore throat' was defined as sore throat or pain or difficulty swallowing *with or without* fever.
- In the Lock 2010 trial, an episode of sore throat was defined as at least three consecutive days with sore throat. Any consecutive recording of sore throat interrupted by four days of non-recording constitutes a new episode.
- Alho 2007 primarily looked for pharyngitis due to group A *Streptococcus* with a positive culture, and secondarily for 'all episodes of pharyngitis'.

- In the Koskenkorva 2013 trial, an episode of sore throat was defined as at least two consecutive days with sore throat.
- Stafford 1986 only reported the proportion of patients who "felt" their symptoms were "cured" after 18 months. This study did not report the number of episodes of sore throat or their severity.

Compared to the more specific categories (a), (b) and (c), the Paradise category (d) - 'all combined' includes a broader spectrum that is more analogous to the van Staaij 2004 category of 'sore throat episode' than 'throat infection'. Thus, in this review, the data on 'episodes of sore throat of any severity' refer in the Paradise studies to the 'all types combined' category (Paradise 1984; Paradise 2002a; Paradise 2002b), and in the van Staaij 2004, Lock 2010 and Koskenkorva 2013 trials to the 'episodes of sore throat' data. For the Alho 2007 study, this refers to the "all episodes of pharyngitis" data rather than "pharyngitis due to group A streptococcus".

We also extracted the more severe sore throat episodes defined as 'moderate or severe' (Paradise 1984; Paradise 2002a; Paradise 2002b) and 'throat infection' (van Staaij 2004). In our analyses these two categories are assumed equivalent; we recorded and analysed these as 'moderate or severe' episodes.

In all the Paradise studies a 'sore throat day' was defined as one on which a sore throat lasted an hour or longer. The comparable data from the van Staaij 2004 and Lock 2010 trials were the number of sore throat days.

Other outcomes

None of the studies stated an intention to record, and nor did they report, consumption of antibiotics or analgesics as outcomes. The studies also did not pre-specify which data on complications or morbidity would be collected and reported, apart from the number of days of pain post-surgery.

Koskenkorva 2013 measured quality of life using the Glasgow Benefit Health Inventory, but this was done only for the tonsillectomy group. van Staaij 2004 used the TACQoL or TAPQoL but did not reported results in a way that could be meta-analysed. Lock 2010 used the PedsQoL.

Excluded studies

After reviewing the full text, we excluded seven studies. Four were non-randomised cohort studies (Clayburgh 2011; Kaiser 1930; Naiboǧlu 2010; Stalfors 2012). We excluded three trials because it was not clear whether the included children had suffered from recurrent acute or chronic tonsillitis, as trial entry criteria were poorly defined (Mawson 1967; McKee 1963; Roydhouse 1970). See Characteristics of excluded studies.

Risk of bias in included studies

Summary assessment of risk of bias

We judged the methodological quality of most included trials to be moderate: see the 'Risk of bias' graph (Figure 2). Details of the risk of bias in the included studies are presented in the 'Risk of bias' summary (Figure 3). We acknowledge that it is impossible for the patients in any tonsillectomy trial to be unaware of their group allocation. The two adult trials were at less risk of bias than the trials involving children, but this was almost entirely related to small losses to follow-up, as a result of short follow-up periods (Alho 2007; Koskenkorva 2013).

Figure 2. 'Risk of bias' graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.
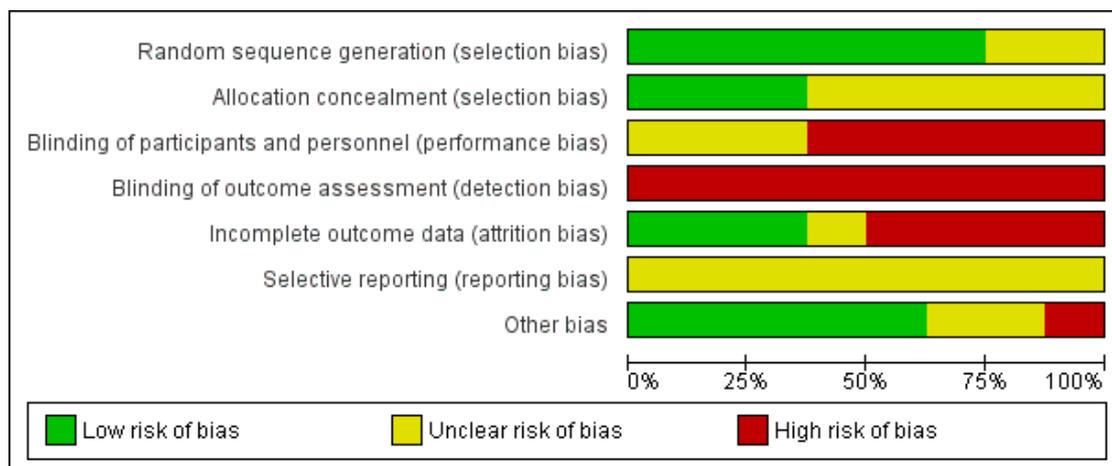
Figure 3. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included study.

Allocation

The method of random sequence generation was adequately described in six of the eight trials, whereas it was unclear in two trials (Paradise 1984; Stafford 1986). Furthermore, concealment of allocation was not adequately described in three (Paradise 1984; Stafford 1986; van Staaij 2004) (see Figure 3 and also the 'Risk of bias' tables in Characteristics of included studies). We had concern about the effectiveness of allocation concealment when blocked randomisation using small blocks of three or four patients was used in these non-blinded studies and there was no description of the methods used to conceal the allocation of existing participants from the team members in charge of recruitment or enrolment of patients (Paradise 2002a; Paradise 2002b).

In one trial baseline characteristics appeared to be imbalanced (Paradise 1984). The children in the surgery groups (tonsillectomy alone or adenotonsillectomy) differed from the control group in terms of the history of episodes of throat infection before entry into the study and in terms of parental socioeconomic status (Paradise 1984, table 1). Those in the surgical group were more often admitted to the trial on the basis of frequent infection in the year prior to entry rather than less frequent infections over a longer period. At the time of our original review we felt that this imbalance may make interpretation of the results problematic (Burton 1999). However, in his response to our review (see 'Feedback' below), Professor Paradise indicates that a number of children met more than one of the frequency-related eligibility criteria (seven or more episodes per year for one year, or five per year for two years, or three per year for three years) and that such children were categorised as meeting the criterion involving the largest number of episodes. It seems that by chance more children in the surgical rather than the control group met the seven plus episodes per year for one year criterion. We stated in the original review that "The surgical group may therefore have included children with more severe disease" and we still believe that this is a possibility. However, as Professor Paradise points out, the resulting bias would favour control subjects and result in a potential underestimation of the treatment effect. We went on to say "Alternatively, these [the children entered because of seven plus episodes in one year] may have been children with less severe, but more short-lived disease in whom a period of frequent infections is more likely to be followed by spontaneous resolution than in those with longer more chronic histories". We accept the observation that some of these children may also have had episodes in previous years, but would continue to argue that our own comment is not counterintuitive. There may be a difference between those children who have a short, relatively brief, period with many sore throats in whom spontaneous reso-

lution then occurs, and those who experience several sore throats per year for many years. A second baseline imbalance was noted. The children in the non-surgical group were more likely to have parents with higher socio-economic status than those in the surgical group. Professor Paradise argues that this difference is likely to favour the control group and again lead to an underestimation of any treatment effect.

Blinding

None of the included trials performed blinding of trial participants due to the nature of the intervention and control treatment. In one trial, researchers who conducted interviews or processed self completed questionnaires/diaries were blind to the interventions of trial participants (Lock 2010). As the main outcomes of the studies were reported by patients (mostly using patient diaries), there was a risk of detection bias.

Incomplete outcome data

All trials reported the numbers of participants who failed to complete the trial and the number of patients excluded.

In seven of the eight included trials intention-to-treat analysis was performed - all patients were analysed according to the group to which they were randomised. In one trial, children in the control group who elected to have surgery and those that did not complete the whole follow-up period were excluded from analysis (Paradise 1984).

In four of the five trials on the effectiveness of adeno-/tonsillectomy in children, we judged the risk of attrition bias to be high due to the large number of participants lost to follow-up (Lock 2010; Paradise 1984; Paradise 2002a; Paradise 2002b). In one study, the trial was stopped at a fixed time period (van Staaij 2004). Although the loss to follow-up in this trial was relatively low, only about half of the children were enrolled at a time point which allowed them to complete a two-year follow-up period.

In the Lock 2010 trial, loss of participants to follow-up at the end of the first and second years was not the only potential source of significant attrition bias. For individual participants, the data sets were incomplete. Data were collected using monthly diaries and self completed postal questionnaires. The mean number of diaries returned per child in the trial was 9.91 out of 24. As a result, it is difficult to work out how the outcome data relating to "mean number of episodes of sore throat per month" or "mean days of sore throat" were calculated. The study reported that the mean episodes and days with sore throat outcomes were estimated by taking the average of the numbers reported per (four-week) diary

returned and multiplying that by 26 to get the rate over two years. It is possible that the diaries were missing in a non-random manner (differential missing rate), introducing a significant risk of bias. The study report noted that there was a possible overestimation of the number of episodes and days with sore throat. When the study authors adjusted for the number of diaries a child returned, the weighted mean number of sore throats decreased over two years from 11.4 to 9.0 for the control group and from 7.4 to 5.5 in the surgery group. We conducted a sensitivity analysis to investigate the impact of inclusion of these trial data in the analysis.

There were a large number of missing data sets in the second and third years of the trials and a large proportion of patients receiving surgery (see Other potential sources of bias), therefore we only included data from the first 12 months of follow-up (Lock 2010; Paradise 1984; Paradise 2002a; Paradise 2002b). The authors of the van Staaij 2004 trial kindly provided us with data on the outcomes at one year.

We judged the risk of attrition bias in the two trials on the effectiveness of tonsillectomy for recurrent pharyngitis in adults to be low (Alho 2007; Koskenkorva 2013; Stafford 1986). There were no losses to follow-up. In one trial, 6% (2/34) of the participants allocated to the control treatment received tonsillectomy during follow-up (Alho 2007), while in the other trial this percentage was 8% (3/40) (Koskenkorva 2013). In the third trial, six out of 20 patients (30%) had received tonsillectomy by the end of the two-year study follow-up (Stafford 1986). In all these trials, all participants allocated to tonsillectomy received surgery.

In studies where patients were analysed according to the groups to which they were randomised, the percentage of patients in the control arm receiving surgery within the first year of randomisation ranged from 6% and 8% (as above) to 15% in the Paradise studies within the first year of follow-up (Paradise 1984; Paradise 2002a; Paradise 2002b). Only the *overall* number of children (over the entire duration of the trial) receiving surgery in the control arm were available for Lock 2010 and van Staaij 2004 and these were 26% (36/137) and 34% (50/149), respectively.

Data on the losses to follow-up and the number of patients in control groups receiving surgery are provided in the 'Risk of bias' tables (Characteristics of included studies).

Selective reporting

We did not have access to study protocols and we did not have enough information to assess the risk of selective reporting bias. Without sufficient information, we judged the risk of bias from selective reporting to be unclear.

As there were only a few studies, we could not conduct formal assessment using funnel plots.

Other potential sources of bias

Baseline characteristics appeared to be balanced in six of the seven trials (Alho 2007; Koskenkorva 2013; Lock 2010; Paradise 2002a; Paradise 2002b; van Staaij 2004).

There is a risk of underestimating the effectiveness of surgery when large numbers of patients in the control group also undergo surgery. This risk appears to be relatively low when the follow-up period is short, but high over periods of two years or more. The proportion in the control group receiving surgery was 6% and 8% for Alho 2007 and Koskenkorva 2013, respectively; they followed up patients for an average of six months. The Paradise studies reported the number of patients receiving surgery at each yearly time point. The cumulative rate was 8% (Paradise 2002a) and 15% (Paradise 1984; Paradise 2002b) over one year, and 11%, 22% and 25% respectively after two years. In the only adult study that collected data for more than a year, the percentage of participants opting for surgery within two years was 30% (Stafford 1986). Lock 2010 and van Staaij 2004 only reported the overall number of children in the control group receiving surgery, i.e. 50/149 (34%) in van Staaij 2004 and 36/137 (26%) in Lock 2010.

Effects of interventions

See: **Summary of findings for the main comparison** Tonsillectomy or adenotonsillectomy versus no surgery in children with chronic/recurrent acute tonsillitis

Although all the studies on the effectiveness of adeno-/tonsillectomy in children extended over several years, large losses to follow-up mean that for the majority of studies sufficient data are only available to allow consideration of results in the first year following surgery. We therefore only included data from the first 12 months of follow-up from the five trials on the effectiveness of adeno-/tonsillectomy in children (Lock 2010; Paradise 1984; Paradise 2002a; Paradise 2002b; van Staaij 2004), and up to six months for the two trials in adults (Alho 2007; Koskenkorva 2013). Stafford 1986 had no losses to follow-up, but did not report any of the outcomes we planned to analyse.

We analysed results in the following way:
- Analysis 1: Tonsillectomy/adenotonsillectomy versus no surgery in children.
- Analysis 2: Tonsillectomy/adenotonsillectomy versus no surgery in adults.
- Analysis 3: Sensitivity analysis of the effect of combining adenotonsillectomy and tonsillectomy versus tonsillectomy alone.
- Analysis 4: Sensitivity analysis based on the risk of bias of included studies.

As baseline severity and frequency of tonsillitis could be factors that affect the results, we analysed the data for children as two subgroups: 'severely affected patients' (Paradise 1984) and 'moderately affected patients' (Lock 2010; Paradise 2002a; Paradise 2002b; van Staaij 2004).

The individual results of the seven trials are summarised in Appendix 3.

Tonsillectomy or adenotonsillectomy versus no surgery

1. Number and severity of episodes of tonsillitis or sore throat

This outcome was measured in two ways in the trials. Some trials only reported episodes of sore throat of <u>any</u> severity, whereas others also reported episodes of 'moderate/severe' sore throat.

Studies conducted in children

Episodes of sore throat of any severity

The first analysis demonstrates the number of episodes of sore throat of any severity <u>including</u> the immediate postoperative period (Analysis 1.1). In this instance we analysed the data using the random-effects model because of the presence of substantial statistical heterogeneity (I^2 = 57%). Surgery is associated with 0.6 fewer sore throat episodes of any severity (95% confidence interval (CI) -1.0 to -0.1 episodes). Children who had an adeno-/tonsillectomy had an average of three episodes of sore throats (of *any* severity) in the first postoperative year, compared to 3.6 episodes in the control group. The 'predicable' episode of pain immediately following surgery is, however, important in interpreting this figure as this accounts for one of the three episodes in the surgical group. We consider the evidence to be of moderate quality, due to lack of standardised methods or definitions in reporting or measuring episodes of sore throat in these studies (which are unblinded) and some uncertainty due to missing data.

Tests for subgroup differences did not show a significant difference between the more severely affected patients and those moderately affected (P value = 0.56, I^2 = 0%). We identified one study as at a high risk of bias for this outcome, due to high proportion of missing data (Lock 2010). We tested the impact of this study in a sensitivity analysis. Removing this study did not have much impact on the effect size (mean difference (MD) -0.42, 95% CI -0.87 to 0.02) or heterogeneity.

Episodes of moderate/severe sore throat

In this analysis, the test for subgroup differences between the more severely affected patients and those moderately affected was statistically significant (P value = 0.0003) (Analysis 1.2), therefore we did not pool the subgroups.

Children who had been more severely affected and had adeno-/tonsillectomy had on average 1.1 episodes of sore throat in the first postoperative year, compared with 1.2 episodes in the control group; a small difference of 0.1 episodes (95% 0.6 fewer to 0.4 more) (Paradise 1984). This is not a significant difference but one episode in the surgical group was that occurring immediately after surgery. There may be a benefit to knowing the precise timing of one 'predictable' episode of pain lasting several days following surgery. We considered the results to be of low quality as they came from one small randomised controlled trial (RCT) and there were more missing data/excluded patients in the control group.

Less severely affected children had more episodes of moderate/severe sore throat after surgery (1.2 episodes) than in the control group (0.4 episodes): a difference of 0.8 episodes (95% CI 0.7 to 0.9 episodes), but again one episode was the postoperative episode. We rated the evidence as moderate quality.

Studies conducted in adults

Episodes of sore throat of any severity

There was substantial statistical heterogeneity in this analysis (I^2 = 94%) (Analysis 2.1) and we analysed the results using the random-effects model. There were 3.6 fewer episodes in the group receiving surgery within six months post-surgery (95% 7.9 fewer to 0.7 more). We rated this evidence as low quality because of the short duration of follow-up (indirectness of evidence) and inconsistency. The difference in effect sizes reported by the two studies is substantial; the reduction was 1.5 fewer episodes (95% CI 2.3 fewer to 0.7 fewer) in Alho 2007 and 5.9 fewer episodes (95% CI 7.8 fewer to 0.7 fewer) in Koskenkorva 2013. When interpreting these findings, it should be noted that the episode of sore throat following surgery has not been included in this analysis.

Episodes of moderate/severe sore throat

The two studies in adults did not distinguish between types of sore throat in ways that were comparable with the studies performed in children.

2. Number of days with sore throat

Studies conducted in children

In the meta-analysis on sore throat days, the number of sore throat days is about five days less in the surgical group (-5.1 days, 95% CI -8.1 to -2.2) (Analysis 1.3). The number of days with pain after surgery has been included in this analysis. Therefore the clinical correlate of this is that rather than having about 23 days of sore

throat in the year, children receiving surgery had 18 days of sore throat, of which some (between five and seven on average) were immediately following surgery and, as such, entirely predictable. There was some statistical heterogeneity (I^2 = 43%), but this was reduced if data from Lock 2010 were excluded from the analysis (then the I^2 = 0%, P value = 0.59). We previously noted that this study had a high risk of bias in overestimating the number of days and episodes of sore throat due to the high percentage of missing data. Removing Lock 2010 from the analysis reduced the mean difference by about one day to 4.3 days (difference -4.3 days, 95% CI -8.0 to -1.3). Due to these concerns, and the absence of blinding, we considered the evidence to be of moderate quality.

Studies conducted in adults

Unlike the studies in children, the studies in adults did not include the number of days post-surgery in their analysis (see Analysis 2.2). There was significant statistical heterogeneity (I^2 = 85%) and we analysed the results using the random-effects model. The pooled mean difference for number of days with sore throat for about six months of follow-up was 10.6 days fewer in the group receiving surgery (95% CI -15.5 to -5.8). The difference in the effect sizes reported by the two studies is substantial; the reduction in mean number of days with pain varied from 8.9 days fewer (95% CI -14.0 to -3.9) in Alho 2007 to 35.1 days fewer (95% CI -54.0 to -16.2) in Koskenkorva 2013. As previously noted, the number of days with sore throat in the immediate postoperative periods is not taken into account in this analysis, which appeared to be 13 (standard deviation (SD) 4) days in Alho 2007 and 17 (SD 6) days in Koskenkorva 2013. In addition, the data in Koskenkorva 2013 are potentially skewed (large standard deviations observed). We considered the evidence for this outcome to be low quality due to these concerns.

3. Morbidity and mortality of surgery

None of the studies listed mortality as an outcome and none reported any deaths in the description of complications arising from surgery.

None of the studies clearly stated which outcomes related to morbidity or complications would be measured in the trials (apart from the number of days with sore throat following surgery). Tonsillectomy is a painful procedure: the days with postoperative pain might reasonably be considered clinically significant. These were already included in the analysis of days of sore throat for children (Analysis 1.3), but not for adults (Analysis 2.2).

For children, the Paradise 1984 trial reports a mean figure for the number of days of pain after surgery as 4.9 days. In the later studies a mean of 6.3 days is given, with a wide range from 0 to 21 days (Paradise 2002a; Paradise 2002b).

In the two trials in adults (aged 13 years and older) (Alho 2007; Koskenkorva 2013), the mean duration of continuous throat pain

following surgery was 13 days (SD 4) and 17 days (SD 6), respectively.

Complications reported in the included studies are summarised in Table 1. In four studies in children primary bleeding was reported in 1.5% to 4.8%, while secondary bleeding was reported in 4.3% and 6% in the two studies in adults.

4. Consumption of antibiotics

None of the studies specifically reported on antibiotic prescriptions or consumption of antibiotics.

5. Consumption of analgesics

Only one study reported data for analgesics consumption (Lock 2010). The mean number of drugs prescribed was two per patient over two years (SD 3.1) in the surgical arm (n = 120) and two (SD 2.6) in the control arm (n = 115).

6. Absence or time off work or school

Studies conducted in children

The results of combining all those children undergoing surgery (tonsillectomy and/or adenotonsillectomy) versus (initial) non-surgical treatment across all studies with follow-up for one year are reported in Analysis 1.4. Children who had surgery had 2.3 fewer days (95% CI -3.4 to -1.2) of absence from school compared to an average of six days of absence among children in the control group. We considered this evidence to be of moderate quality. However, the days off school immediately following surgery are not included in this analysis (and data in '3. Morbidity and mortality of surgery' above indicate this period to amount to five to six days on average).

Studies conducted in adults

Only Koskenkorva 2013 reported on this outcome. Absenteeism was 3.3 days lower in the surgery arm, but this was not significant (95% CI -7.7 to 1.1) (Analysis 2.3).

7. Quality of life

Three studies collected quality of life data, but these unfortunately could not be pooled (Koskenkorva 2013; Lock 2010; van Staaij 2004)

Lock 2010 collected data at three months, 12 months and 24 months using the parent form of PedsQoL (Paediatric Quality of Life Inventory). The amount of missing data at 12 months was significant; only 64% of participants (71/111) in the surgery arm and 48% of participants (52/108) in the control arm returned the PedsQoL questionnaire.

van Staaij 2004 collected data for health-related quality of life using the TNO-AZL (Netherlands Organisation for Applied Scientific

Research Academic Medical Centre) Preschool Children's Quality Of Life Questionnaire (TAPQoL) for children aged two to five or the TNO-AZL Children's Quality Of Life Questionnaire (TAC-QoL) for children aged five years and older and the Child Heath Questionnaire - Parental Form (CHQ pf50). The study reported that small differences were found in some domains of the quality of life questionnaires at six months but these were "not clinically relevant". No differences were found in the other domains or at 24 months.

Koskenkorva 2013 used a post-intervention questionnaire (Glasgow Benefit Inventory, GBI), which was only applied to the intervention group.

The quality of life data are summarised below.

Study	Quality of life instrument	Scale	Score
Lock 2010[a]	PedsQoL	Physical health	Surgery: 90.0 (SD 16.4), n = 71 Control: 85.3 (SD 17.9), n = 52
		Psychosocial health	Surgery: 83.8 (SD 15.3), n = 71 Control: 80.0 (SD 17.5), n = 52
van Staaij 2004	Arcola or TAPQoL CHQ - parental form	All subscales	"No difference" at 24 months
Koskenkorva 2013[b]	Glasgow Benefit Inventory (GBI) (Scale range from -100 (maximal negative benefit) to 0 (no benefit) to +100 (maximal benefit) Robinson 1996) Data were only collected from the tonsillectomy group	Overall score	27 (SD 12)
		General health subscore	23 (SD 15)
		Social subscore	3 (SD 12)

[a] Lock 2010: table here shows the 12-month data.
[b] Koskenkorva 2013: data were only available for the intervention group. The GBI is a generic post-intervention questionnaire that assesses the effects of surgery on the health status of patients.

Tonsillectomy versus adenotonsillectomy

Data from the Paradise 2002a study allowed direct comparison between tonsillectomy and adenotonsillectomy. No significant differences were found in either the number of episodes of sore throat or sore throat days. As reported above, in an earlier trial results from those children undergoing tonsillectomy and adenotonsillectomy were combined because no differences between the groups were observed (Paradise 1984).

For this update we undertook an appropriate sensitivity analysis by analysing these as subgroups. We did not observe any heterogeneity in treatment effects between different types of surgery (Analysis 3.1; Analysis 3.2; Analysis 3.3), but this should be interpreted with caution as this analysis may not be powered to detect small differences, and it is possible that such differences could be masked by other forms of clinical heterogeneity present between studies. Overall, we may never know whether the absence of evidence of a difference between the effects of the procedures is because no such difference exists, or is due to a lack of statistical power to detect a difference. There remains doubt about whether or not removing the adenoids (reducing or preventing as it does the incidence of mouth breathing) has an effect on the frequency and/or severity of sore throats.

Effectiveness of surgery after the first postoperative year

Four of the five included trials suffered from substantial loss to follow-up, especially during the second and third year of the trials (Lock 2010; Paradise 1984; Paradise 2002a; Paradise 2002b). As a consequence, good information was only available for the first 12 months.

In the previous version of this review we presented data from the van Staaij 2004 trial and suggested that it might be possible to draw some tentative conclusions from these data about the effect of tonsillectomy in the second postoperative year. However, having clarified the methods underlying the data (see the individual results section in Appendix 3), it is not possible to determine what happened to participants specifically in the second year of follow-up. The study does include data from 139 participants who were followed up for 23 months, of whom 80 participants were followed for two years. However, data on the frequency and duration of sore throats during that second year alone are not available. The proportion of the 174 participants who spent 23 months in the trial, who were lost to follow-up, is 20% and therefore similar to the other studies.

Stafford 1986 only reported the number of patients who "felt" their symptoms were "cured" after 18 months. This is neither a predefined outcome in our analysis protocol, nor one that could be pooled with other studies. Between 18 months and two years after enrollment, 18/20 participants in the surgery group versus 14/20 participants in the antibiotics group felt that their symptoms were "cured" (risk ratio 0.86, 95% CI 0.57 to 1.29).

DISCUSSION

Summary of main results

• This review includes seven trials with low to moderate risk of bias: five undertaken in children (987 participants) and two in adults (156 participants). Good information about the effectiveness of adeno-/tonsillectomy is only available for the first year following surgery in children and for a shorter period (five to six months) in adults

• We combined data from five trials in children. Those children who had an adeno-/tonsillectomy had an average of three episodes of sore throats (of *any severity*) in the first postoperative year, compared to 3.6 episodes in the control group; a difference of 0.6 episodes (95% confidence interval (CI) -1 to -0.1; moderate quality evidence). One of the three episodes in the surgical group was the 'predictable' one that occurred in the immediate postoperative period.

• When we analysed only episodes of *moderate/severe* sore throat, **children who had been more *severely affected*** and had

adeno-/tonsillectomy had on average 1.1 episodes of sore throat in the first postoperative year, compared with 1.2 episodes in the control group (low quality evidence). This is not a significant difference but one episode in the surgical group was that occurring immediately after surgery.

• **Less severely affected children** had more episodes of *moderate/severe* sore throat after surgery (1.2 episodes) than in the control group (0.4 episodes: difference 0.8, 95% CI 0.7 to 0.9), but again one episode was the predictable postoperative episode (moderate quality evidence).

• Data on the number of sore throat days are only available for **moderately affected children** and are consistent with the data on episodes. In the first year after surgery children undergoing surgery had an average of 18 days of sore throat (of which some - between five and seven on average - will be in the immediate postoperative period), compared with 23 days in the control group (difference 5.1 days, 95% CI 2.2 to 8.1; moderate quality evidence).

• Data for **adults** were only available for analysis for up to five to six months of follow-up. When we pooled data from two studies (156 participants), there were 3.6 fewer episodes (95% CI 7.9 fewer to 0.70 more; low quality evidence) in the group receiving surgery within six months post-surgery. However, statistical heterogeneity was significant. The pooled mean difference for number of days with sore throat in a follow-up period of about six months was 10.6 days fewer in favour of the group receiving surgery (95% CI 5.8 fewer to 15.8 fewer; low quality evidence). Again there was significant statistical heterogeneity in this analysis and the number of days with postoperative pain (which appeared to be on average 13 to 17 days in the two trials) was not included. Given the short duration of follow-up and the differences between studies, we considered the evidence from adults to be of low quality.

• Two studies in children reported that there was "no statistically significant difference" in quality of life outcomes but the data could not be pooled. One study reported no difference in analgesics consumption. We found no evidence for prescription of antibiotics.

Overall completeness and applicability of evidence

In evaluating the data from the available trials of adeno-/tonsillectomy for chronic or recurrent acute tonsillitis or sore throats there are several issues that must be considered.

• Only one study examines the effectiveness of adeno-/tonsillectomy in severely affected children (children with the so-called 'Paradise criteria' for surgery).

• Most trials included children with more modest symptoms and disease.

- It is not possible to determine whether or not the effectiveness of tonsillectomy and adenotonsillectomy are different.
- Adequate data are only available to evaluate the effectiveness of adeno-/tonsillectomy in children in the first postoperative year due to substantial loss to follow-up in the included trials, especially in the second and third years.
- Three trials assessed the effectiveness of tonsillectomy in adults, but only data from the short-term studies (duration of follow-up five to six months) could be pooled. One trial determined the short-term effectiveness of tonsillectomy in one specific subgroup of patients (recurrent episodes of pharyngitis due to group A *Streptococcus*), while the other trial reported the short-term effectiveness of tonsillectomy in patients with recurrent episodes of pharyngitis of any origin. The trial that reported data up to 24 months was small (40 participants), at high risk of bias and only reported the proportion of patients who "felt cured".
- There are no trials evaluating the intermediate to long-term effectiveness of tonsillectomy in adults (other than one at high risk of bias).

While an adult or child who has had their palatine tonsils removed cannot suffer from tonsillitis they can suffer from sore throats. Unfortunately, the palatine tonsils are just one part of Waldeyer's ring of lymphoid tissue within the pharynx. The aggregates of lymphoid tissue in the tongue base, nasopharynx and within the pharyngeal walls, along with the remaining soft tissues of the pharynx, remain after tonsillectomy. This, at least in part, explains why tonsillectomy is not a panacea for all sore throats. Patients with chronic or recurrent acute sore throats are heterogeneous in terms of both causation and severity. Some have symptoms primarily due to infection of the tonsils, others do not. Thus an important issue is the role the palatine tonsils play in producing the clinical condition of 'tonsillitis', or indeed 'sore throat' or 'throat infection'. It is self evident that removal of the palatine tonsils will prevent 'tonsillitis' just as removing the appendix will prevent appendicitis. However, many patients with 'tonsillitis' will have inflammation, or infection, of other pharyngeal lymphoid tissue and other soft tissues of this region. If inflammation/infection of these tissues is critically dependent on inflammation/infection of the tonsils themselves, tonsillectomy might be effective. If inflammation/infection of non-tonsillar tissue is independent of tonsillar infection, removing the tonsils may be irrelevant.

Children

The studies provide limited information on the effectiveness of adeno-/tonsillectomy in certain types of children. While some of the children probably suffered from a clinical condition in which the palatine tonsils played a major (perhaps exclusive) role in the pathogenesis of their symptoms and signs, others might have been included in whom the palatine tonsils were not the source of their symptoms. In his initial study, Paradise tried extremely hard to identify children who fell into the first group (Paradise 1984). As a result the 'Paradise criteria' have become widely recognised as a standard set of criteria for children with severe throat symptoms due to tonsillar disease in whom tonsillectomy is being considered. Equally, when following up children post-randomisation, Paradise had strict criteria for episodes of throat infection. The later Paradise studies specifically looked at less severely affected children and it is certainly possible that this group included children in whom the palatine tonsils played a less important role in their symptoms (Paradise 2002a; Paradise 2002b). The van Staaij 2004 trial is unique because the eligibility criteria were much broader. Any child whose ENT surgeon felt they should undergo adeno-/tonsillectomy was potentially eligible. Those fulfilling the 'Paradise criteria' were excluded, leaving as the included group children who might have been similar to the Paradise (2002) 'less severely affected' children (Paradise 2002a; Paradise 2002b). However, equally, some children may have fulfilled even less stringent criteria. To reiterate, these were children who ENT surgeons working in 21 general hospitals and three academic centres in the Netherlands had deemed fulfilled their normal criteria for surgery. Whilst specific information on the nature and severity of sore throats prior to randomisation is lacking in the van Staaij 2004 trial, no such criticism can be made of the follow-up period. A sore throat was defined as sore throat or pain or difficulty in swallowing combined, with or without a fever. The study by Lock 2010 initially included both children experiencing mild symptoms (six or more episodes within two years or eight or more episodes within one year) and moderate symptoms (sore throat for five days or more: six or more episodes within two years or six or more within one year). These stringent criteria were, however, amended midway through the study and the inclusion criteria were changed to include children experiencing four or more episodes of sore throat within each of two years or six or more episodes of sore throat within one year. As a consequence the entry criteria for inclusion in this trial were less stringent than those of Paradise 1984.

In summary, what the included studies actually evaluate is the effectiveness of adeno-/tonsillectomy for the symptoms experienced by three more or less heterogeneous groups of children, and these studies themselves are heterogeneous.

- In Paradise 1984 many (possibly most) included children had symptoms related to their palatine tonsils so this group is likely to be least heterogeneous.
- Paradise 2002a, Paradise 2002b and Lock 2010 had a heterogeneous group of less severely affected children. What proportion had symptoms related to their palatine tonsils?
- van Staaij 2004 had potentially the most heterogeneous group of less severely affected children. Again, the proportion with symptoms related to their palatine tonsils is uncertain.

Is it, therefore, rational to combine the studies given the clinical heterogeneity among the included study populations of the trials? Combining results would probably lead to an underestimate of a

positive treatment effect from surgery in the most severely affected children and an overestimate in the least severely affected.

Severely affected children

Is there a positive treatment effect in the most severely affected children? These children from Paradise 1984 are included in our first meta-analysis (Analysis 1.1; Analysis 1.2; Analysis 1.3). In terms of moderate/severe sore throat, the benefit of adeno-/tonsillectomy is exchanging the 1.2 episodes experienced by the control group for 1.1 episodes following surgery at a 'cost' of the one episode following surgery, which means the children had only an average of 0.1 fewer episodes. The postoperative episode might be thought to be preferable because it can be timed and anticipated. However, this is at best a modest benefit and, given the width of the confidence interval (95% CI 0.6 fewer to 0.4 more), the true value may be even less. The commensurate reduction in any type of sore throat was 0.9 episodes (3.1 versus 2.2 (= 1.2 + 1 episode post-surgery)). No significant reduction was found in terms of number of sore throat days in the first year (18.9 in the control group versus 16.3 in the adeno-/tonsillectomy group) but (on average) five to seven of these 16.3 days were in the immediate postoperative period and hence highly predictable. Since there are no 'post-surgical days' in later years, one might speculate that had sufficient data been available for the second and third years post-surgery, a difference may have been seen.

Less severely affected children

As might be expected, the benefits seen in the less severely affected children are less (Analysis 1.1; Analysis 1.2; Analysis 1.3), with a increase in *moderate or severe* sore throat episodes of 0.8, and a decrease of *any* sore throat of 0.5, if the episode of pain due to surgery is taken into account. In this analysis there was a statistically significant reduction in sore throat days from 24 to 18 days. That period of 18 days includes the predictable five to seven days of pain in the immediate postoperative period.

Effects in the second year post-surgery

Little useful information is available about the effects of surgery in the second postoperative year and beyond.

Adults

Two trials examined the short-term effectiveness of tonsillectomy in adults (Alho 2007; Koskenkorva 2013). One trial focused on a specific subgroup of patients (recurrent episodes of pharyngitis due to group A *Streptococcus*), while the other trial included patients with recurrent episodes of pharyngitis from any origin. Although both studies suggest that tonsillectomy offers short-term benefits as compared to (initial) non-surgical treatment, the limited duration of follow-up prevents us from drawing firm conclusions. We must conclude that we do not know, one way or the other, whether tonsillectomy is generally helpful for adults suffering from sore throats in the intermediate or long term. If we could be certain that (a) the symptoms were solely attributable to inflammation of their tonsils, and (b) these symptoms were going to continue, then tonsillectomy would be a reasonable treatment option. It is difficult (if not impossible) to be certain that both of these conditions are true for an individual patient. Although a third trial reported percentages of people who "felt cured", the trial had many limitations and the effect size observed was not significant.

Quality of the evidence

We judged the evidence for the primary outcomes for children's data to be of moderate quality; we downgraded the evidence from high quality mainly due to the high number of losses to follow-up that introduced a substantial risk of (attrition) bias, especially beyond one year after surgery, and also due to inconsistencies in the data collection.

Most of the evidence is derived from trials in children and there is a lack of data for adults, particularly after six months of follow-up. There was also significant heterogeneity observed in the data for adults. Therefore, the quality of evidence for the main outcomes for adults is low.

Potential biases in the review process

The outcome definitions and methods of data collection were different between studies, introducing some clinical heterogeneity. However, we did standardise the definition of key outcomes such as (refraining from) including the number of days and episodes of sore throat post-surgery in the meta-analyses. There was also a concern about potential skewness from the use of count data for episodes of sore throat.

Agreements and disagreements with other studies or reviews

Our main results are in agreement with findings reported in a recent (narrative) systematic review on the effects of tonsillectomy for recurrent sore throats in children (Barraclough 2014), and with current clinical practice guidelines on tonsillectomy for children and adults with recurrent sore throats (Baugh 2011; SIGN 2010). The information about complications provided by the studies included in this review is consistent with other published data even though reporting may be incomplete. Relatively small randomised controlled trials are often inadequate for providing accurate data about rare complications. The best available data on primary and secondary haemorrhage following tonsillectomy are provided by the United Kingdom National Tonsillectomy Audit (Brown 2005); the primary and secondary haemorrhage rates were

0.6% and 3%, respectively, in a group of 33,921 patients who underwent tonsillectomy over a 14-month period in 2003 and 2004.

AUTHORS' CONCLUSIONS

Implications for practice

Good information about the effectiveness of adeno-/tonsillectomy is only available for the first year following surgery in children and a shorter period (five to six months following surgery) in adults.

For more **severely affected children** (those who fulfil the 'Paradise criteria'), those who have an adeno-/tonsillectomy will have a significant, but predictable episode of sore throat immediately following surgery and then two other episodes in the first postoperative year (1 predictable + 2 unpredictable = 3 in total) compared with 3.6 unpredictable episodes in the control group. The difference is 0.6 episodes, but with a confidence interval of -1 to -0.1; in other words, the real difference may be one episode or almost none. In practical terms, this difference is modest.

Less severely affected children who have adeno-/tonsillectomy may never have had another *moderate/severe* sore throat anyway; the chance of them so doing is modestly reduced by adeno-/tonsillectomy. For them, surgery will mean having an average of 1.2 moderate/severe sore throats (1 predictable + 0.2 unpredictable = 1.2 in total) rather than 0.4 unpredictable episodes experienced by the control group. If the postoperative episode is set aside (and this may not be a reasonable thing to do), there is no difference in the number of unpredictable episodes of moderate/severe sore throat. The children who have surgery will have 18 rather than 23 sore throat days but something between five to seven (on average) of these 18 days will be in the immediate postoperative period.

Current evidence suggests that in adults with recurrent episodes of pharyngitis tonsillectomy may not lead to a clinically significant reduction in the number of sore throat episodes and sore throat days in the first five to six months following surgery as compared to (initial) non-surgical treatment.

The very modest benefits that have been identified should be carefully weighed against the possible harms of surgery, as the procedure is associated with a small but significant degree of morbidity in the form of primary and secondary haemorrhage and, even with good analgesia, is particularly uncomfortable for adults.

Questions related to the natural history of recurrent/chronic tonsillitis/'sore throat'

One reason why the impact of surgery is so modest is that many patients in the control group get better spontaneously. There is a trade-off for the physician and patient, who must weigh up a number of different uncertainties. These are:

Uncertainty 1: Are my symptoms primarily attributable to disease of my palatine tonsils?

Uncertainty 2: Am I going to continue to have sore throats and, if so, will they be as bad, better or worse?

Decision: Do I want to exchange the uncertainties of both the possibility and timing of my sore throats, in the year ahead, for the certainty of a predictable period of postoperative pain that may be followed by fewer sore throats than I might have had?

Many thousands of patients have undergone adeno-/tonsillectomy and are likely to continue to do so. Some will undergo this surgery 'unnecessarily' because they would not have had any further throat problems had they not had surgery. During a period in the 1980s and 1990s in the United Kingdom, when waiting lists for surgery were long, a proportion of patients who were listed for surgery according to strict criteria eventually declined the operation because they had improved spontaneously. Perhaps the price society pays for the prevention of recurrent sore throats in one group of patients is that for each of these individuals an indeterminable number of other individuals will have their tonsils removed 'unnecessarily'.

Those who choose surgery for themselves or their child must be fully informed of the risks of the procedure; this is critical in conducting the appropriate harm-benefit analysis. Evidence from the included studies confirms that primary and secondary haemorrhages are still significant complications. Even more robust data are provided by the United Kingdom National Tonsillectomy Audit, undertaken between 2003 and 2004. The audit collected information from 33,921 patients undergoing tonsillectomy (Brown 2005). Overall the rates of primary and secondary haemorrhage were 0.6% and 3%, respectively; the rates varied according to the surgical technique used to remove the tonsils.

Those considering tonsillectomy or adenotonsillectomy for themselves or their children, as a treatment for chronic or recurrent acute tonsillitis, and those advising them, should be aware of two important uncertainties that may affect their treatment decisions. They must acknowledge some uncertainty about whether or not their symptoms are primarily due to their tonsils and realise that adeno-/tonsillectomy is not a panacea for all types of sore throat. There is also uncertainty about the likelihood that these will continue in the future, which is only partly predictable from the frequency and severity of symptoms they have experienced in the past.

The benefits of surgery are greatest in children fulfilling the 'Paradise criteria'. In their original setting these were very strict, requiring a child to have had each of the following: (1) seven or more episodes in the preceding year, or five or more in each of the preceding two years, or three or more in each of the preceding three years, comprising (2) characteristic clinical features, having been (3) treated with antibiotics when streptococcal infection was proven or suspected, and (4) each episode had to have been be

documented in a clinical record. In day-to-day practice these criteria are often 'watered down', with most emphasis being placed on the aspect of 'counting' the number of episodes, and less on proving the 'certainty' of each episode.

Surgery is associated with a reduction in the number of unpredictable days with sore throat and the number of episodes of sore throat even in those who are less severely affected. One of the key practical issues is that the episode of a period of days with sore throat that follows surgery is entirely predictable, lasting between about five and seven days on average, but in some cases being shorter and in others considerably longer. These benefits should be carefully weighed against the possible harms of surgery as adeno-/tonsillectomy is associated with a small but significant degree of morbidity in the form of primary and secondary haemorrhage and, even with good analgesia, is particularly uncomfortable for adults. It may have to be accepted that some patients will undergo 'unnecessary surgery' in order for others to benefit, and that even well-informed and appropriately counselled patients are prepared to accept the risks and uncertainty of surgery.

Implications for research

Evidence

One implication drawn from the trials is that greater benefit can potentially be obtained by more severely affected children. If the 'Paradise' criteria are accepted for identifying such children, what are the equivalent criteria for adults? If the most severely affected adults could be identified, would a randomised trial, looking at outcomes over several years, be appropriate or practicable?

Population

Adults with recurrent tonsillitis.

Intervention

Tonsillectomy or adenotonsillectomy.

Comparison

Usual care, watchful waiting or other medical management.

Outcomes

Episodes of, and number of days with, sore throat and a measure of their severity, number of days of absence from school or work, quality of life, complications of surgery, such as haemorrhage, and number of days with pain postoperatively.
Data should be reported in a way that allows the episode and pain due to surgery to be taken into account.

Timing

At least one year and preferably two or three years.
There would be little need to consider further studies to evaluate the effectiveness of adeno-/tonsillectomy if three things were certain. Firstly, that an individual's throat problems were entirely due to their tonsils, secondly that these problems were bound to disappear if they had surgery, and finally that they would definitely continue if they did not have surgery. There would certainly be no need for a randomised trial. Two appropriate research questions therefore might be:

1. Is it possible to identify those individuals in whom the presence of the palatine tonsils is a critical determinant of their clinical condition?

2. Is it possible to better predict which individuals will continue to have problems and to predict the severity of those problems?

It is unlikely to be easy, and may not be possible, to address these uncertainties.

REFERENCES

References to studies included in this review

Alho 2007 *{published data only}*

Alho O-P, Koivunen P, Penna T, Teppo H, Koskela M, Luotonen J. Tonsillectomy versus watchful waiting in recurrent streptococcal pharyngitis in adults: randomised controlled trial. *BMJ* 1997;**334**(7600):939. [PUBMED: 17347187]

Koskenkorva 2013 *{published data only}*

Koskenkorva T, Koivunen P, Koskela M, Niemela O, Kristo A, Alho OP. Short-term outcomes of tonsillectomy in adult patients with recurrent pharyngitis: a randomized controlled trial. *CMAJ: Canadian Medical Association Journal* 2013; **185**(8):E331–6. [DOI: 10.1503/cmaj.121852; PUBMED: 23549975

Lock 2010 *{published data only}*

* Lock C, Wilson J, Steen N, Eccles M, Mason H, Carrie S, et al. North of England and Scotland Study of Tonsillectomy and Adeno-tonsillectomy in Children (NESSTAC): a pragmatic randomised controlled trial with a parallel non-randomised preference study. *Health Technology Assessment* 2010;**14**(13):1-164, iii-iv. [DOI: 10.3310/hta14130

Wilson JA, Steen IN, Lock CA, Eccles MP, Carrie S, Clarke R, et al. Tonsillectomy: a cost-effective option for childhood sore throat? Further analysis of a randomized controlled trial. *Otolaryngology - Head and Neck Surgery* 2010;**146**:122–8.

Paradise 1984 *{published data only}*

Paradise JL, Bluestone CD, Bachman RZ, Colborn DK, Bernard BS, Taylor FH, et al. Efficacy of tonsillectomy for recurrent throat infection in severely affected children. *New England Journal of Medicine* 1984;**310**(11):674–83. [PUBMED: 6700642]

Paradise 2002a *{published data only}*

Paradise JL, Bluestone CD, Colborn DK, Bernard BS, Rockette HE, Kurs-Lasky M. Tonsillectomy and adenotonsillectomy for recurrent throat infection in moderately affected children. *Pediatrics* 2002;**110**(1 (Pt 1)): 7–15. [PUBMED: 12093941]

Paradise 2002b *{published data only}*

Paradise JL, Bluestone CD, Colborn DK, Bernard BS, Rockette HE, Kurs-Lasky M. Tonsillectomy and adenotonsillectomy for recurrent throat infection in moderately affected children. *Pediatrics* 2002;**110**(1 (Pt 1)): 7–15. [PUBMED: 12093941]

Stafford 1986 *{published data only}*

Stafford N, von Haacke N, Sene A, Croft C. The treatment of recurrent tonsillitis in adults. *Journal of Laryngology and Otology* 1986;**100**(2):175–7. [PUBMED: 3950480]

van Staaij 2004 *{published and unpublished data}*

van Staaij BK, van den Akker EH, Rovers MM, Hordijk GJ, Hoes AW, Schilder AG. Effectiveness of adenotonsillectomy in children with mild symptoms of throat infections or adenotonsillar hypertrophy: open, randomised controlled trial. *Clinical Otolaryngology* 2005;**30**(1):60–3. [PUBMED: 15748193]

* van Staaij BK, van den Akker EH, Rovers MM, Hordijk GJ, Hoes AW, Schilder AGM. Effectiveness of adenotonsillectomy in children with mild symptoms of throat infections or adenotonsillar hypertrophy: open, randomised controlled trial. *BMJ* 2004;**329**(7467):651. [PUBMED: 15361407]

References to studies excluded from this review

Clayburgh 2011 *{published data only}*

Clayburgh D, Milczuk H, Gorsek S, Sinden N, Bowman K, MacArthur C. Efficacy of tonsillectomy for pediatric patients with dysphagia and tonsillar hypertrophy. *Archives of Otolaryngology -- Head and Neck Surgery* 2011;**137**: 1197–202.

Kaiser 1930 *{published data only}*

Kaiser AD. Results of tonsillectomy. A comparative study of twenty-two hundred tonsillectomized children with an equal number of controls three and ten years after operation. *Journal of the American Medical Association* 1930; **95**:837–42.

Mawson 1967 *{published data only}*

* Mawson SR, Adlington P, Evans M. A controlled study evaluation of adeno-tonsillectomy in children. *Journal of Laryngology and Otology* 1967;**81**:777–90.

Mawson SR, Adlington P, Evans M. A controlled study evaluation of adeno-tonsillectomy in children. *Journal of Laryngology and Otology* 1967;**82**:963–79.

McKee 1963 *{published data only}*

McKee WJE. A controlled study of the effects of tonsillectomy and adenoidectomy in children. *British Journal of Preventive and Social Medicine* 1963;**17**:49–69.

Naiboǧlu 2010 *{published data only}*

Naiboǧlu B, Külekçi S, Kalayɑ k C, Sheidaei S, Toros SZ, Egeli E. Improvement in quality of life by adenotonsillectomy in children with adenotonsillar disease. *Clinical Otolaryngology* 2010;**35**:383–9.

Roydhouse 1970 *{published data only}*

Roydhouse N. A controlled study of adenotonsillectomy. *Archives of Otolaryngology* 1970;**92**:611–6.

Stalfors 2012 *{published data only}*

Stalfors J, Ericsson E, Hemlin C, Hultcrantz E, Månsson I, Roos K, et al. Tonsil surgery efficiently relieves symptoms: analysis of 54 696 patients in the National Tonsil Surgery Register in Sweden. *Acta Oto-Laryngologica* 2012;**132**: 533–9.

References to ongoing studies

NATTINA 2014 *{published data only}*

* HTA - 12/146/06: The NAtional Trial of Tonsillectomy IN Adults (NATTINA): a clinical and cost effectiveness study. http://www.nets.nihr.ac.uk/projects/hta/1214606.

Additional references

Ashworth 2004

Ashworth M, Latinovic R, Charlton J, Cox K, Rowlands G, Gulliford M. Why has antibiotic prescribing for respiratory illness declined in primary care? A longitudinal study using the General Practice Research Database. *Journal of Public Health (Oxford)* 2004;**26**:268-74.

Barraclough 2014

Barraclough J, Anari S. Tonsillectomy for recurrent sore throats in children: indications, outcomes, and efficacy. *Otolaryngology - Head and Neck Surgery* 2014;**150**:722–9.

Baugh 2011

Baugh RF, Archer SM, Mitchell RB, Rosenfeld RM, Amin R, Burns JJ, et al. Clinical practice guideline: tonsillectomy in children. *Otolaryngology - Head and Neck Surgery* 2011; **144**:S1–30.

Blackshaw 2014

Blackshaw H, Zhang L-Y, Venekamp RP, Wang B, Chandrasekharan D, Schilder AGM. Tonsillectomy versus tonsillotomy for sleep-disordered breathing in children. *Cochrane Database of Systematic Reviews* 2014, Issue 11. [DOI: 10.1002/14651858.CD011365

Brown 2005

Brown P, Ryan R, Yung M, et al. National Prospective Tonsillectomy Audit. http://www.rcseng.ac.uk/publications/docs/national_prospective.html/view?searchterm=tonsillectomy%20audit 2005:1–40.

Handbook 2011

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions version 5.1.0 [updated March2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Marshall 1998

Marshall T. A review of tonsillectomy for recurrent throat infection. *British Journal of General Practice* 1988;**48**: 1331–5.

Paradise 1996

Paradise JL. Tonsillectomy and adenoidectomy. In: Bluestone CD, Stool SE, Kenna MA editor(s). *Pediatric Otolaryngology*. 3rd Edition. Philadelphia: WB Saunders, 1996:1054–65.

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Robinson 1996

Robinson K, Gatehouse S, Browning GG. Measuring patient benefit from otorhinolaryngological surgery and therapy. *Annals of Otology, Rhinology and Laryngology* 1996; **105**:415–22.

SIGN 2010

Scottish Intercollegiate Guidelines Network - Guideline 117 - Management of sore throat and indications for tonsillectomy - April 2010. www.sign.ac.uk/pdf/sign117.pdf.

Venekamp 2014

Venekamp RP, Hearne BJ, Chandrasekharan D, Blackshaw H, Lim J, Schilder AGM. Tonsillectomy or adenotonsillectomy versus non-surgical management for sleep-disordered breathing in children. *Cochrane Database of Systematic Reviews* 2014, Issue 6. [DOI: 10.1002/14651858.CD011165

References to other published versions of this review

Burton 1999

Burton MJ, Towler B, Glasziou P. Tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis. *Cochrane Database of Systematic Reviews* 1999, Issue 3. [DOI: 10.1002/14651858.CD001802

Burton 2009

Burton MJ, Glasziou PP. Tonsillectomy or adeno-tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis. *Cochrane Database of Systematic Reviews* 2009, Issue 1. [DOI: 10.1002/14651858.CD001802.pub2

* *Indicates the major publication for the study*

Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Review)   **28**

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Alho 2007

Methods	**Allocation:** randomised - replacement randomisation **Design:** parallel groups, average follow-up about 6 months
Participants	**Number:** 70 adults out of 298 screened **Age:** 15 years and above **Setting:** ENT Department of the University of Oulu Hospital (Finland) **Eligibility criteria:** • 3 or more episodes of pharyngitis in 6 months or 4 episodes in 12 months • The signs and symptoms during the episodes had to be typical of streptococcal pharyngitis. In addition, these episodes had to be severe enough for the patient to seek medical attention and at least 1 episode had to be group A streptococcal infection proved by culture or rapid antigen test **Exclusion criteria:** • History of peritonsillar abscess, ongoing antibiotic treatment for other illness, recurrence probably caused by non-compliance with treatment, major heart or airway disorder or bleeding diatheses that would make same day surgery unfeasible, and residence outside the city of Oulu or the neighbouring 8 communities
Interventions	**Intervention group:** tonsillectomy (total extracapsular tonsillectomy using blunt or diathermy dissection); n = 36 **Comparator group:** watchful waiting (remaining on waiting list); n = 34 **Use of additional interventions:** participants in the control group did not receive prophylactic treatment for their tonsillitis. No information was provided on any standardised treatment strategy in case of tonsillitis recurrence during follow-up
Outcomes	**Primary outcome:** Proportion of patients with an acute episode of group A streptococcal pharyngitis during the 90 days' follow-up, as determined by signs and symptoms of acute pharyngitis with a positive result of throat culture **Secondary outcomes:** • Percentage change in the proportion of patients with all episodes of pharyngitis at 90 days • The times to episodes during the whole follow-up period • *The difference in the mean rates of episodes during the whole follow-up period* • *Days with symptoms (fever, sore throat, rhinitis, cough) during the whole follow-up period* • *All episodes of pharyngitis (an episode was considered to be at least 2 consecutive days with sore throat)* The mean length of follow-up was 164 (SD 63) in the control group and 170 (SD 12) in the tonsillectomy group Patients recorded episodes and days with symptoms in diaries All participants were advised to visit their own general practitioner if they had acute symptoms suggestive of pharyngitis. The general practitioner would then take a culture sample from the pharynx and send it to Oulu University Hospital for analysis. Patients

	were given written instructions for their general practitioner about the study and how to obtain the culture sample (from surface of both tonsils or tonsillar fossae in patients who had undergone tonsillectomy and the posterior pharyngeal wall). The patients were told that it was important to seek medical advice for their symptoms during the trial in exactly the same way they had done before the trial and that it was possible to have streptococcal pharyngitis after tonsillectomy
Funding sources	None declared
Declarations of interest	None declared
Notes	Baseline characteristics were balanced

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quotes: "generated the randomisation sequence with a computer random number generator". "To avoid disparity between group sizes, we used replacement randomisation"
Allocation concealment (selection bias)	Low risk	Quotes: "... concealed from the investigators who enrolled the participants ... sequentially numbered sealed opaque envelopes. These were opened sequentially only after an eligible participant had been found and informed consent obtained..." "This task was performed by research assistant not involved in the assignment of care to the trial participants"
Blinding of participants and personnel (performance bias) All outcomes	High risk	No blinding. No information about standardised treatment for sore throats during follow-up
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding. Outcomes were patient-reported (subjective)
Incomplete outcome data (attrition bias) All outcomes	Low risk	No loss to follow-up at 90 days. Patients analysed according to group randomised
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit a judgement of low or high risk

Other bias	Low risk	No other important sources of bias were found
		Proportion of patients who did not receive assigned treatment:
		• Participants in intervention group who did not have surgery: 0/36 (0%)
		• Participants in control group who had surgical treatment within 90 days: 2/ 34 (5%), 1 due to pain, another to coincide with another sibling undergoing surgery

Koskenkorva 2013

Methods	**Allocation:** randomised - simple randomisation
	Design: parallel groups, average follow-up about 6 months

Participants	**Number:** 86 adults out of 260 screened
	Age: 13 years and above
	Setting: tertiary care ENT centre in Oulu (Finland)
	Eligibility criteria:
	• 3 or more episodes of pharyngitis within the previous 12 months
	• These episodes had to be disabling, prevent normal functioning, be severe enough for the patient to seek medical attention and be thought to involve the palatine tonsils
	• It was not necessary for culture or antigen tests to have shown infection with group A *Streptococcus*
	Exclusion criteria:
	• History of peritonsillar abscess, chronic tonsillitis, ongoing antibiotic treatment, residence outside of the Oulu region, pregnancy or previous illness making same-day surgery unfeasible

Interventions	**Intervention group:** tonsillectomy (total extracapsular tonsillectomy using blunt or diathermy dissection); n = 46 (45 patients underwent tonsillectomy and 1 patient underwent adenotonsillectomy)
	Comparator group: watchful waiting (placement on waiting list for tonsillectomy to undergo surgery after 5 to 6 months); n = 40
	Use of additional interventions: no information was provided on a standardised treatment strategy in cases where there was an acute sore throat/pharyngitis recurrence during follow-up

Outcomes	**Primary outcome:**
	• Proportion of patients who had a severe episode of pharyngitis within 5 months. A severe episode was defined as acute throat pain and signs suggesting the symptoms originated in the pharynx (oedema, erythema, exudative tonsillitis, anterior cervical lymphadenitis), which led to medical consultation. In addition, the serum level of C-reactive protein had to be higher than 40 mg/L either at the day of the medical consultation or 3 days later. If a blood sample was not taken, the result of a throat culture had to show other than normal flora, and the patient had to grade the throat

	pain as severe
	Secondary outcomes:
	• Proportion of patients with any episode of pharyngitis (sore throat lasting at least 2 days) and episodes with medical consultation during the 5-month follow-up
	• Times to pharyngitis episodes
	• *Reported number of sore throat episodes*
	• *Reported number of sore throat days*
	• *Reported number of days absent from school or work*
	• Reported number of symptomatic days during follow-up
	• *Health-related quality of life* (Glasgow Benefit Inventory) in the intervention group (Robinson 1996)
	• *Adverse effects related to tonsillectomy*
	The patients used a symptom diary to record the presence and severity (mild, moderate or severe) of throat pain, cough, rhinitis, fever and absence from school or work. Data concerning acute visits and tonsillectomy were recorded from patients' charts
	Patients were advised to visit the study physician or their general practitioner whenever they had acute symptoms suggestive of pharyngitis. In addition, they were told that it was important to seek medical advice for their symptoms during the trial exactly as they had done before. At the acute visit, patients underwent a clinical examination including a throat swab and a blood test to measure serum levels of C-reactive protein. The blood test was repeated 3 days later. All laboratory and microbiological analyses were performed by staff blinded to the clinical data

Funding sources	None declared
Declarations of interest	None declared
Notes	Baseline characteristics were balanced

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "We assigned patients ... using simple randomisation"
Allocation concealment (selection bias)	Low risk	Quote: "The allocation sequence was concealed from the investigators using sequentially numbered sealed opaque envelopes"
Blinding of participants and personnel (performance bias) All outcomes	High risk	No blinding. No additional information about management of sore throats
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding. Outcomes were patient-reported (subjective)

Incomplete outcome data (attrition bias) All outcomes	Low risk	No losses to follow-up 2 patients, 1 in each group, lost their symptom diaries but reported not having consulted a physician for throat pain. These patients were assumed to have no symptoms during the study period
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit a judgement of low or high risk
Other bias	Low risk	No other sources of bias were found **Proportion of patients who did not receive assigned treatment:** • Participants in intervention group who did not have surgery: 0/46 (0%) • Participants in control group who had surgical treatment within 5 months: 3/40 (8%)

Lock 2010

Methods	**Allocation:** randomised. Stratification for age (3 age groups) and study centre with use of balanced blocks **Design:** parallel groups. 2-year follow-up. Study was done alongside a parallel non-randomised preference study
Participants	**Number:** 268 children (729 patients enrolled into study out of 1546 screened; the other children were enrolled into the non-randomised arms of the study) **Age:** 4 to 15 years **Setting:** 5 secondary care ENT departments located in the north of England or west of Scotland (UK) **Eligibility criteria:** At the start of the trial (May 2002) the inclusion criteria were: • children (or carers) reporting experience of mild symptoms, 6 or more episodes within 2 years or 8 or more episodes within 1 year, *or* • children reporting experience of moderate symptoms (sore throat for 5 days or more), 6 or more episodes within 2 years or 6 or more within 1 year Mid-way through the study (May 2004), the inclusion criteria were amended to: • children (or carers) reporting experience of 4 or more episodes of sore throat within each of 2 years or 6 or more episodes of sore throat within 1 year **Exclusion criteria:** • Previous tonsillectomy, hospitalisation due to tonsillitis, quinsy, marked obstructive airway during attack • Comorbidity affecting ability to undergo surgery within 6 months • Bleeding disorder, otitis media with effusion, sleep apnoea syndrome, rare medical condition (e.g. glomerulonephritis or Henoch-Schönlein purpura), suspected velopharyngeal insufficiency, congenital/valvular heart disease

Interventions	**Intervention group:** tonsillectomy (dissection or bipolar diathermy according to surgical preference) or adenotonsillectomy (tonsillectomy with adenoid curettage); n = 131 **Comparator group:** standard non-surgical (medical) treatment; n = 137 **Use of additional interventions:** the referring general practitioner was free to treat recurrences of sore throat according to his or her standard practice
Outcomes	**Primary outcome:** • Reported number of episodes of sore throat in the 2 years after entry into the study (these data were presented as the mean number of episodes per month in Year 1 and Year 2) **Secondary outcomes:** • Reported number of episodes of sore throat, otitis media and upper respiratory tract infection that invoked a general practitioner consultation • Reported severity of sore throats • Reported number of symptom-free days • Quality of life (PedsQoL, parent-reported) • Surgical and anaesthetic morbidity Data on sore throats were gathered by a simple structured daily health diary, which was completed and returned by participants on a monthly basis for 24 months (monthly diaries). In addition, outcome questionnaires were administered to study participants at 3, 12 and 24 months into the study. Data on consultation rates and prescribed medication were gathered from general practitioner's medical records by manual abstraction by trained researchers at the end of follow-up for all trial participants The study aimed to follow up all participants for 24 months from the date of initial randomisation; this was not done
Funding sources	UK Department of Health through the National Institute of Health Research, Health Technology Assessment Programme
Declarations of interest	None declared
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Independent ... web-based computer randomisation"
Allocation concealment (selection bias)	Low risk	Quote: "Independent ... web-based computer randomisation". Blocked randomisation was used; unclear how large these blocks were
Blinding of participants and personnel (performance bias) All outcomes	High risk	No blinding. The referring general practitioner was free to treat recurrences of sore throat according to his or her standard practice

Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "... all researchers who conducted interviews or processed self-completed questionnaires and diaries were blind to interventions of all ... participants" Comment: since the content of questionnaires and diaries was based on the reports of participants/carers who were not blind, it is not clear how effective the described strategy would be in reducing detection bias. Furthermore, the main outcomes, such as number of episodes and days with sore throat, were estimated based on patient diaries and were subjective. The question asked was "Was your child bothered by any of the following today?" and "sore throat" was an option

Incomplete outcome data (attrition bias)
All outcomes
	High risk	There was a very high risk of bias from the loss of data and imputations
		Significant loss of outcome data because of the following:
		1. Loss of participants (not included in analyses); the reasons for loss to follow-up not described
		• **Participants not included in analyses at Year 1:** 37/268 (14%)
		    ○ Intervention arm: 12/131 (9%)
		    ○ Control arm: 25/137 (18%)
		• **Participants not included in analyses at Year 2:** 111/268 (41%)
		    ○ Intervention arm: 48/131 (37%)
		    ○ Control arm: 63/137 (46%)
		2. Loss of outcome data from monthly diaries (mean number of diaries returned per participant was 9.91 out of a maximum of 24 (41%) over 2 years)
		Data on number of episodes and days with sore throat were collected through diaries. Given the loss of participants and outcome data, it was unclear how the following was estimated:
		• "mean number of episodes of sore throat per month" at end of Year 1 and Year 2 (Table 12 of Lock 2010)
		• "unadjusted estimate" of "mean number of sore throats" over 2 years (Table 16 of Lock 2010)

Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Review)   **35**

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lock 2010 *(Continued)*

| | | • "mean days of sore throat in Year 1 and Year 2" (Table 17 of Lock 2010) The number of episodes and days with sore throat pain over 2 years seemed to be estimated by taking the mean number of sore throat per (4-week) diary returned multiplied by 26. However, the numbers in Table 17 of Lock 2010 do not seem to be completely consistent with the application of this method

• The average number of diaries returned per child was 9.91, out of 24 4-weekly diaries sent (page 11). The proportion of diaries returned per child was low (9.91/24 = 41%)

• Imputations were done for missing data for the non-returned diaries, but the data could have been missing in a non-random manner; e.g. patients may be more likely to return a diary that has been filled in when they were symptomatic

• The study report noted that there was a possible overestimation of the number of episodes and days with sore throat. When adjusted for the number of diaries a child returned, the weighted mean number of sore throats decreased from 11.4 to 9.0 in the medical group and 7.4 to 5.5 in the surgery group over 2 years

3. Loss of outcome data from questionnaires at 12 and 24 months for quality of life data
Participants for whom outcome questionnaire data were available at 12 months - intervention arm:
Year 1 - 73/131 (58%)
Year 2 - 64/131 (49%)
Participants for whom outcome questionnaire data were available at 12 months - comparator arm:
Year 1 - 55/137 (40%)
Year 2 - 54/137 (39%)
Participants were analysed by group randomised |
| Selective reporting (reporting bias) | Unclear risk | Insufficient information to judge whether low risk or high risk of bias |

Other bias	Unclear risk	Subjects were analysed by group randomised (ITT), but one-quarter of patients in the control arm received surgical intervention **Proportion of patients who did not receive assigned treatment at the end of 2 years (Year 1 data not available)** • Participants in intervention group who did not have surgery: 11/131 (8%) • Participants in control group who had surgical treatment: 36/137 (26%)

Paradise 1984

Methods	**Allocation:** randomised; stratification for age (3 age groups) with use of balanced blocks of 4 **Design:** parallel groups. Study was done alongside a non-randomised cohort study
Participants	**Number:** 91 children randomised out of 2043 patients screened **Age:** 3 to 15 years **Setting:** Children's Hospital of Pittsburgh (USA). Participants were recruited among hospital outpatients, children referred by community practitioners or directly by parents between August 1971 and June 1982 **Eligibility criteria:** • History of recurrent episodes of throat infection (i.e. tonsillitis, pharyngitis or tonsillopharyngitis) The episodes had to meet defined standards in each of 4 categories: (1) Frequency of occurrence: 7 or more episodes in the preceding year, 5 or more in each of the 2 preceding years, or 3 or more in each of the preceding 3 years (2) Clinical features: each episode characterised by 1 or more of the following: oral temperature of at least 38.3 °C, cervical lymphadenopathy (enlarged (> 2 cm) or tender cervical lymph nodes), tonsillar or pharyngeal exudate, or positive culture for group A beta-haemolytic *Streptococcus* (3) Treatment: antibiotics administered at conventional dosage for proven or suspected streptococcal episodes (4) Documentation: each episode and its qualifying features substantiated by concurrent notation in a medical record **Exclusion criteria:** • Judged to require prompt removal of large tonsils or adenoids because of proven or suspected alveolar hypoventilation, sustained difficulty in swallowing or marked discomfort in breathing • Prior tonsil or adenoid surgery, major physical or emotional disease, structural middle-ear damage or sensorineural hearing loss, hypogammaglobulinaemia (IgG, IgA or IgM level > 2 SD below the mean value for the child's age) • Simultaneous enrolment of a sibling in the study, or inability to meet the projected schedule of follow-up visits

Interventions	**Intervention group 1:** tonsillectomy (dissection and snare technique); n = 27 **Intervention group 2:** adenotonsillectomy (dissection and snare technique for tonsillectomy; reverse adenotomes and curettes for adenotomy); n = 16 **Comparator group 1:** non-surgical treatment; n = 29 **Comparator group 2:** non-surgical treatment; n = 19 The trial data were pooled (and not presented separately) in the publication: **Intervention group:** tonsillectomy or adenotonsillectomy; n = 43 **Comparator group:** non-surgical treatment; n = 48 **Use of additional interventions:** cultures for the presence of group A streptococci in case of suspected or diagnosed throat infections showing growth of any degree were considered positive. Penicillin V (250 mg) - or erythromycin (10 mg/kg body weight) in case of presumed allergy to penicillin - was prescribed 4 times daily for 10 days for all participants with positive cultures and also for those whose cultures were negative but who had been treated presumptively from the outset and had improved markedly within 48 hours

Outcomes	**Primary outcome:**
	• Observed episodes of throat infections based on 3 criteria:
	  ○ moderate/severe sore throat;
	  ○ group A streptococcal infection; *or*
	  ○ "counting" method.
	Secondary outcomes:
	• Proportion of visits with isolated cervical lymphadenopathy
	• Number of parent-reported sore throat days
	• Number of days of sore throat-associated school absence
	A standardised telephone or in-person inquiry was made bi-weekly concerning the day-by-day occurrence of specified symptoms and events, including sore throat and absence from school. Standardised clinical examinations were performed at 6-week intervals and at the time of respiratory illnesses

Funding sources	ER McCluskey Memorial Research. Programme fund from the Children's Hospital in Pittsburgh and National Institute of Child Health and Human Development
Declarations of interest	None declared
Notes	Not balanced for history of episodes of throat infection before trial entry and parental socioeconomic status (2 of 11 comparisons): 65% in the control group were skilled workers or executives/professionals, whereas 65% in the surgery group were of semi-skilled/unskilled workers or disabled/unemployed/receiving public assistance/other. 47% in the surgery group had more than 7 infections within the past year, whereas only 23% in the control group met this criterion. 67% in the control group met the criterion of >= 3 infections per year over the past 3 years

Risk of bias

Bias	Authors' judgement	Support for judgement

Random sequence generation (selection bias)	Unclear risk	Quote: "assigned randomly, within categories and balanced blocks of four"
Allocation concealment (selection bias)	Unclear risk	Comment: not described. Small block size (4 participants) potentially allows prior knowledge of allocation
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No blinding. However, there was a clear protocol for management of reported sore throats
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding. Patients were followed up bi-weekly by either a standardised telephone or in-person inquiry. Patients reported their symptoms
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: significant number of patients not included in analyses especially in years 2 and 3. Patients were excluded from analysis for these reasons 1) loss to follow-up; 2) not treated as randomised (patients in the control arm who received surgery were excluded from analysis); 3) did not complete the full follow-up periods in the analysis Proportion of patients not included in analysis is higher in the control groups Children electing for surgery in the control group **Participants not included in analyses total:** Year 1 - 18/91 (20%) ○ Intervention arm: 5/43 (12%) ○ Control arm: 13/48 (27%) Year 2 - 31/91 (34%) ○ Intervention arm: 12/43 (28%) ○ Control arm: 19/48 (40%) Year 3 - 49/91 (54%) ○ Intervention arm: 21/43 (49%) ○ Control arm: 28/48 (58%)
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit a judgement of low or high risk
Other bias	Unclear risk	**Proportion of patients who did not receive assigned treatment:** • Participants in intervention group who did not have surgery within 90 days: 8/43 (19%) • Participants in control group who

		had surgical treatment (cumulative numbers)
		     ○ Year 1 = 7/48 (15%)
		     ○ Year 2 = 12/48 (25%)
		     ○ Year 3 = 16/48 (33%)
		Baseline characteristics: not balanced for history of episodes of throat infection before trial entry and parental socioeconomic status (2 of 11 comparisons)

Paradise 2002a

Methods	**Allocation:** randomised; stratification for age (3 age groups) and history (whether the frequency or clinical features of previous history met the eligibility criteria for the Paradise 1984 trial) with use of balanced blocks of 3 **Design:** parallel groups, 3-armed trial. Study was done alongside a 2-arm study (see Paradise 2002b)

Participants	**Number:** 177 children (373 patients were eligible for either of the studies, Paradise 2002a or Paradise 2002b, from 2174 patients screened)
	Age: 3 to 15 years
	Setting: Children's Hospital of Pittsburgh (USA). Participants were recruited among hospital outpatients, children referred by community practitioners or directly by parents between April 1982 to March 1994
	Eligibility criteria:
	• History of recurrent episodes of throat infection (i.e. tonsillitis, pharyngitis or tonsillopharyngitis). The episodes had to meet slightly less stringently defined standards regarding either the frequency of previous episodes or their clinical features or their degree of documentation than those used in the Paradise 1984 trial. Details are described in full in the appendix of the publication, but in summary these include:
	• ○ Aged 3 to 15 years: 7 or more counting episodes in the preceding year, 5 or more counting episodes in each of the 2 preceding years, or 3 or more counting episodes in each of the preceding 3 years. If all previous episodes were undocumented, 1 qualifying unit observed by a study team clinician within 4 months after initial observation. If previous episodes were partially documented, notation in a clinical record indicating the occurrence of 1 qualifying unit within the 3 months period preceding initial evaluation
	    ○ Aged 3 to 6 years: 5 or 6 qualifying units in the preceding year, including at least 2 counting episodes or 4 qualifying units in each of the past 2 years including at least 2 counting episodes in each year
	    ○ Aged 7 to 15 years: 4 to 6 qualifying units in the preceding year, including at least 1 counting episode or 3 qualifying units in each of the past 2 years including at least 1 counting episode in each year
	• No apparent indications for adenoidectomy (obstructing adenoids or history of recurrent or persistent otitis media)
	Exclusion criteria:
	• Children judged to require prompt removal of large tonsils or adenoids because of obstructive symptoms

	• Children who met the eligibility criteria for the earlier tonsillectomy trials ([Paradise 1984](#))
Interventions	**Intervention group 1:** tonsillectomy (dissection and snare technique); n = 58 **Intervention group 2:** adenotonsillectomy (dissection and snare technique for tonsillectomy; reverse adenotomes and curettes for adenotomy); n = 59 **Comparator group:** non-surgical treatment; n = 60 **Use of additional interventions:** cultures for the presence of group A *Streptococcus* in case of suspected or diagnosed throat infections showing growth of any degree were considered positive. Penicillin V (250 mg) - or erythromycin (10 mg/kg body weight) in case of presumed allergy to penicillin - was prescribed 4 times daily for 10 days for all patients with positive cultures and also for those whose cultures were negative but who had been treated presumptively from the outset and had improved markedly within 48 hours

Outcomes	**Primary outcome:**
	• Observed episodes of throat infections based on 3 criteria:
	  ○ moderate/severe sore throat;
	  ○ group A streptococcal infection; *or*
	  ○ "counting" method.
	Secondary outcomes:
	• Proportion of visits with isolated cervical lymphadenopathy
	• Number of parent-reported sore throat days
	• Number of days of sore throat-associated school absence
	A standardised telephone or in-person inquiry was made bi-weekly concerning the day-by-day occurrence of specified symptoms and events, including sore throat and absence from school. Standardised clinical examinations were performed at 6-week intervals and at the time of respiratory illnesses

Funding sources	National Institute of Child Health and Human Development. The Children's Hospital of Pittsburgh waived insurance charges and professional fees for patients who lacked insurance coverage
Declarations of interest	None declared
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "... using computer generated random number lists"
Allocation concealment (selection bias)	Unclear risk	Quote: "Assignments were made by designated non-clinical staff members using separate, computer generated random number lists" Small block size (3 participants) potentially

		allows prior knowledge of allocation. There was no blinding and it was unclear whether the person(s) making the decisions about enrolment were aware of the preceding allocations
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No blinding. However, there was a clear protocol for management of sore throat
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding. Outcomes were patient-reported (subjective)

Incomplete outcome data (attrition bias)
All outcomes
	High risk	Comment: a significant number of participants were not included in analyses due to loss to follow-up, especially in years 2 and 3
		Participants not included in analyses total:
		Year 1 - 30/177 (17%)
		    ○ Intervention group 1: 11/58 (19%)
		    ○ Intervention group 2: 13/59 (22%)
		    ○ Control group: 6/60 (10%)
		Year 2 - 53/177 (30%)
		    ○ Intervention group 1: 17/58 (29%)
		    ○ Intervention group 2: 23/59 (39%)
		    ○ Control group: 13/60 (22%)
		Year 3 - 73/177 (41%)
		    ○ Intervention group 1: 22/58 (38%)
		    ○ Intervention group 2: 31/59 (53%)
		    ○ Control group: 20/60 (33%)
		Reasons for loss to follow-up not described, and it was unclear if those lost to follow-up were also patients who did not receive intervention as randomised
		Subjects were analysed based on the groups randomised (ITT analysis). No data imputation for loss to follow-up

Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit a judgement of low or high risk

Other bias	Low risk	**Proportion of patients who did not receive assigned treatment:**
		• Participants in intervention group 1 who did not have surgery: 5/58 (9%)
		• Participants in intervention group 2 who did not have surgery: 5/59 (8%)
		• Participants in control group who had surgical treatment (cumulative numbers)
		   ○ Year 1 - 5/60 (8%)
		   ○ Year 2 - 7/60 (12%)
		   ○ Year 3 - 12/60 (20%)
		Baseline characteristics: balanced

Paradise 2002b

Methods	**Allocation:** randomised; stratification for age (3 age groups) and history (whether the frequency or clinical features of previous history met the eligibility criteria for the Paradise 1984 trial) in balanced blocks of 4 **Design:** parallel-group. This study was done alongside a 3-arm study (see Paradise 2002a)

Participants	**Number:** 151 children (373 patients were eligible for either of the study, Paradise 2002a or Paradise 2002b from 2174 patients screened)
	Age: 3 to 15 years
	Setting: Children's Hospital of Pittsburgh (USA). Participants were recruited among hospital outpatients, children referred by community practitioners or directly by parents between April 1982 to March 1994
	Eligibility criteria:
	• History of recurrent episodes of throat infection (i.e. tonsillitis, pharyngitis or tonsillopharyngitis). The episodes had to meet slightly less stringently defined standards regarding either the frequency of previous episodes or their clinical features or their degree of documentation than those used in the Paradise 1984 trial. Details are described in full in the appendix of the publication, but in summary these include:
	• ○ Aged 3 to 15 years: 7 of more counting episodes in the preceding year, 5 or more counting episodes in each of the 2 preceding years, or 3 or more counting episodes in each of the preceding 3 years. If all previous episodes were undocumented, 1 qualifying unit observed by a study team clinician within 4 months after initial observation. If previous episodes were partially documented, notation in a clinical record indicating the occurrence of 1 qualifying unit within the 3 months period preceding initial evaluation
	   ○ Aged 3 to 6 years: 5 or 6 qualifying units in the preceding year, including at least 2 counting episodes or 4 qualifying units in each of the past 2 years including at least 2 counting episodes in each year
	   ○ Aged 7 to 15 years: 4 to 6 qualifying units in the preceding year, including at least 1 counting episodes or 3 qualifying units in each of the past 2 years including at least 1 counting episodes in each year
	• One or more indications for adenoidectomy (obstructing adenoids or histories of recurrent or persistent otitis media)

	Exclusion criteria: • Children judged to require prompt removal of large tonsils or adenoids because of obstructive symptoms • Children who met the eligibility criteria for the earlier tonsillectomy trial (Paradise 1984)
Interventions	**Intervention group:** adenotonsillectomy (dissection and snare technique for tonsillectomy; reverse adenotomes and curettes for adenotomy); n = 73 **Comparator group:** non-surgical treatment; n = 78 **Use of additional interventions:** cultures for the presence of group A *Streptococcus* in case of suspected or diagnosed throat infections showing growth of any degree were considered positive. Penicillin V (250 mg) - or erythromycin (10 mg/kg body weight) in case of presumed allergy to penicillin - was prescribed 4 times daily for 10 days for all patients with positive cultures and also for those whose cultures were negative but who had been treated presumptively from the outset and had improved markedly within 48 hours
Outcomes	As in Paradise 2002a
Funding sources	National Institute of Child Health and Human Development. The Children's Hospital of Pittsburgh waived insurance charges and professional fees for patients who lacked insurance coverage
Declarations of interest	None declared
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "... using computer generated random number lists"
Allocation concealment (selection bias)	Unclear risk	Quote: "Assignments were made by designated non-clinical staff members using separate, computer generated random number lists." Small block size (4 participants) potentially allows prior knowledge of allocation. There was no blinding and it was unclear whether the person(s) making the decisions about enrolment were aware of the preceding allocations
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No blinding. However, there was a clear protocol for the management of sore throats

Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding. Outcomes were patient-reported (subjective)

Incomplete outcome data (attrition bias)
All outcomes
	High risk	Significant number of participants not included in analyses due to loss to follow-up
		Participants not included in analyses total:
		• Year 1 - 25/151 (17%)
		• ○ Intervention group: 14/73 (19%)
		    ○ Control group: 11/78 (14%)
		• Year 2 - 39/151 (26%)
		    ○ Intervention group: 23/73 (32%)
		    ○ Control group: 16/78 (21%)
		• Year 3 - 53/151 (35%)
		    ○ Intervention group: 29/73 (40%)
		    ○ Control group: 23/78 (29%)
		Reasons for loss to follow-up not described and it was unclear if those lost to follow-up were also patients who did not receive the intervention as randomised
		Participants were analysed based on the groups randomised (ITT analysis). No data imputation for loss to follow-up

Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit a judgement of low or high risk

Other bias	Low risk	**Proportion of patients who did not receive assigned treatment:**
		• Participants in intervention group 1 who did not have surgery: 8/73 (11%)
		• Participants in control group who had surgical treatment (cumulative numbers)
		    ○ Year 1 - 12/78 (15%)
		    ○ Year 2 - 17/78 (22%)
		    ○ Year 3 - 19/78 (24%)
		Baseline characteristics: balanced

Stafford 1986

Methods	**Allocation:** randomised **Design:** prospective, parallel

Participants	**Number:** 40
	Age: range 16 to 42 years
	Setting: ENT outpatient department of tertiary centre in London. Patients were referred by general practitioners
	Male/female: 3/17 in surgery group, 6/14 in antibiotic group
	Average number of tonsillitis episodes per year: 6 in tonsillectomy, 5 in antibiotics group
	Eligibility criteria:
	  • 16 years or older
	  • A history of at least 4 episodes of tonsillitis per year for at least 2 years; each episode consisting of sore throat, dysphagia, pyrexia and general malaise, and lasting a minimum of 3 days
	    • Normal full blood count and negative Paul Bunnell test

Interventions	**Intervention group:** tonsillectomy; n = 20 **Comparator group:** antibiotics, n = 20 Penicillin V 250 mg, 4 times per day for 8 days for each episode of tonsillitis for the following year Patients with penicillin allergy are given either co-trimoxazole "2 tablets" 2 times daily or erythromycin 250 mg 4 times per day **Use of additional interventions:** none described
Outcomes	Time point of measure unclear, most likely between 18 months to 2 years • "Cure rate" as reported by patients Patients in the surgery group were reviewed at 6, 12 and 18 months postoperatively Patients in the antibiotics group were reviewed every 3 months in outpatients. GPs were sent questionnaires to complete and return each time a patient was seen and prescribed an antibiotic for tonsillitis
Funding sources	No information provided
Declarations of interest	No information provided
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "randomly allocated" No description of allocation concealment method
Allocation concealment (selection bias)	Unclear risk	No description of allocation concealment method

Blinding of participants and personnel (performance bias) All outcomes	High risk	No blinding. Patients, investigators and general practitioners in charge of follow-up of both groups were aware of treatment allocation
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: no blinding. Outcomes were patient-reported (whether patients "felt cured" of symptoms)
Incomplete outcome data (attrition bias) All outcomes	Low risk	All patients analysed
Selective reporting (reporting bias)	Unclear risk	No access to protocol. Insufficient information to judge
Other bias	High risk	Follow-up frequency and setting were different between groups. Method of eliciting results most likely different. 6/20 patients in the control arm had received tonsillectomy within 2 years

van Staaij 2004

Methods	**Allocation:** randomised; computer-generated random list of 4 numbers in each block and fixed blocks within each hospital **Design:** parallel-group. Patients were followed up to 2 years, or the end of the trial, whichever was earlier

Participants	**Number:** 300 children out of 1226 screened
	Age: 2 to 8 years
	Setting: ENT surgeons from 21 general hospitals and 3 academic centres in The Netherlands
	Eligibility criteria:
	  • Adenotonsillectomy indicated according to "current medical practice" in the Netherlands. ENT surgeons were asked to provide the indication they considered most important for surgery: recurrent throat infections (3 or more a year) or other indications such as obstructive problems or recurrent upper respiratory tract infections
	Exclusion criteria:
	  • Children meeting strict Paradise 1984 criteria
	  • Children with suspected obstructive sleep apnoea
	  • Down's syndrome, craniofacial malformations such as cleft palate, and immunodeficiency other than deficiencies of IgA or IgG

Interventions	**Intervention group:** adenotonsillectomy; n = 151 **Comparator group:** watchful waiting; n = 149 **Use of additional interventions:** the parents, general practitioners and ENT surgeons were encouraged to manage sore throats and upper respiratory tract infections during follow-up in line with their regular practice

Outcomes	**Primary outcome:**
	• Incidence of fever (a temperature of 38.0 °C or higher) for at least 1 day
	Secondary outcomes:
	• Throat infections
	• Sore throat
	• Upper respiratory tract infection
	• Absence from day-care or school due to upper respiratory tract infection
	• Health-related quality of life
	• Patterns of sleep and eating
	• Height and weight.
	Note: authors were contacted to obtain 1-year data
	The parents kept a diary of upper respiratory tract infections in their child, which included sore throat, pain or difficulty in swallowing, cough, rhinorrhoea, earache and otorrhoea and absences from day-care or school due to upper respiratory tract infections. Parents measured the child's temperature daily with a validated tympanic membrane thermometer
	During scheduled follow-up visits at 3, 6, 12, 18 and 24 months, health-related quality of life questionnaires were completed and ENT examination was carried out
Funding sources	Dutch Healthcare Insurance Board
Declarations of interest	None declared
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "randomisation was by a computer generated list of 4 numbers in each block and fixed blocks within each hospital"
Allocation concealment (selection bias)	Unclear risk	Comment: not described. Small block size (4 participants) potentially allows prior knowledge of allocation
Blinding of participants and personnel (performance bias) All outcomes	High risk	No blinding. No further information about management of sore throats
Blinding of outcome assessment (detection bias) All outcomes	High risk	Although the primary outcome measure of the study was "fever", measured objectively and with the results stored automatically, the outcomes of interest in this review may have been subject to detection bias

van Staaij 2004 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	The study was completed at a fixed time point (February 2003), and the length of follow-up for patients varied, depending on the time of recruitment Loss to follow-up relatively low: **Participants not included in analyses total:** 43/300 (14%); reasons for loss to follow-up described **Participants not included in analyses - intervention group:** 18/151 (12%) **Participants not included in analyses - comparator group:** 25/149 (17%) The amount of data available for each time point was the following (data from authors): For 12 months (denominator indicated the potential number of patients available; i.e. who could have completed the follow-up for the length of time): • Total number of children available in analysis: 262 • Number of children allocated to ATE and completed 12 months: 111/131 (84.7%) • Number of children allocated to WW and completed 12 months: 107/131 (81.7%) For 23 months: • Total number of children available in analysis: 174 • Number of children allocated to ATE and completed 24 months: 70/88 (79.5%) • Number of children allocated to WW and completed 24 months: 69/86 (80.2%) For 24 months: • Total number of children available in analysis: 166 • Number of children allocated to ATE and completed 24 months: 42/83 (50.6%) • Number of children allocated to WW and completed 24 months: 38/83 (45.8%)
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit a judgement of low or high risk

Other bias	Low risk	**Proportion of patients who did not receive assigned treatment:** • Participants in intervention group who did not receive surgery: 7/151 (5%) • Participants in comparator group who received surgical treatment: 50/149 (34%) **Baseline characteristics** - balanced

ATE: adenotonsillectomy
ITT: intention-to-treat
SD: standard deviation
WW: watchful waiting

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Clayburgh 2011	**Allocation**: non-randomised, prospective cohort study
Kaiser 1930	**Allocation:** non-randomised, retrospective cohort study
Mawson 1967	**Participants:** recurrent adenitis alone considered an indication for inclusion; some participants had no attacks in year prior to trial or number was unknown (Mawson 1967, Table VI)
McKee 1963	**Allocation:** randomised using hospital number; quasi-randomisation **Participants:** children (< 15 years) with throat infections or "acute upper respiratory infections with cervical adenitis"
Naiboǧlu 2010	**Allocation:** non-randomised, prospective cohort study
Roydhouse 1970	**Allocation**: "Selection of cases [as] described by McKee" (randomised using hospital number); quasi-randomisation **Participants**: children (2 to 13 years) with "recurrent ... tonsillitis and other respiratory tract infections"
Stalfors 2012	**Allocation:** non-randomised, prospective cohort study

Characteristics of ongoing studies *[ordered by study ID]*

NATTINA 2014

Trial name or title	The NAtional Trial of Tonsillectomy IN Adults (NATTINA): a clinical and cost effectiveness study
Methods	A mixed methods study incorporating a feasibility study, internal pilot and a phase III, multicentre trial randomising patients to immediate tonsillectomy versus conservative management. Patients will be randomised on a 1:1 basis
Participants	Adults (16 years +) referred to secondary care because of disabling recurrent sore throats
Interventions	Intervention: immediate tonsillectomy - dissection of the palatine tonsils Comparison: conservative management, i.e. deferred surgery

Outcomes	**Primary outcomes:**
	   1. Total number of sore throat days over the 24 months following randomisation
	 Secondary outcomes:
	   1. Outcomes measured using NATTINA Sore Throat Alert Reports (STARs) over 24 months
	     i) Number of sore throats in the category (a) mild or (b) moderate/severe
	     ii) Use of over-the-counter medication
	     iii) Use of any prescription therapies
	     iv) Number of days unable to undertake usual activities
	     v) SF-12 general health questionnaire referring to the episode
	   2. Tonsil Outcome Inventory 14 (TOI14) total and subscales (throat, general health, resources and social/psychological) 6-monthly
	   3. SF-12 general health questionnaire 6-monthly
	   4. Nature of any adverse events experienced
	   5. Evaluation of the impact of alternative sore throat patient pathways by observation and statistical modelling of outcomes against baseline severity, as assessed by total number of sore throat days recalled in the 6 months prior to recruitment plus TOI14 score at recruitment.
	   6. Economic evaluation has 3 components (1) a cost-effectiveness analysis, based on the incremental cost per sore throat day avoided (2) cost-utility analysis, based on incremental cost per quality-adjusted life year (QALY) gained (3) cost-benefit analysis, based on patients willingness to pay to avoid a sore throat day

Starting date	April 2014. Expected publication July 2019
Contact information	Chief investigator: Prof Janet Wilson Trial manager: Isabel Rubie (isabel.rubie@ncl.ac.uk) (+44) 0191 2087252
Notes	Details from HTA website: http://www.nets.nihr.ac.uk/projects/hta/1214606

DATA AND ANALYSES

Comparison 1. Tonsillectomy or adenotonsillectomy versus no surgery in children

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Episodes of sore throat of any severity (including as one episode the period post-surgery) at 12 months	5	795	Mean Difference (IV, Random, 95% CI)	-0.56 [-1.04, -0.07]
1.1 Paediatric patients - severely affected patients	1	73	Mean Difference (IV, Random, 95% CI)	-0.90 [-1.92, 0.12]
1.2 Paediatric patients - moderately affected patients	4	722	Mean Difference (IV, Random, 95% CI)	-0.51 [-1.07, 0.05]
2 Episodes of moderate/severe sore throat (including as one episode the period post-surgery) at 12 months	4		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.1 Paediatric patients - severely affected patients	1	73	Mean Difference (IV, Random, 95% CI)	-0.09 [-0.57, 0.39]
2.2 Paediatric patients - moderately affected patients	3	491	Mean Difference (IV, Random, 95% CI)	0.82 [0.70, 0.94]
3 Sore throat days (including those immediately post-surgery) at 12 months	5	776	Mean Difference (IV, Fixed, 95% CI)	-5.13 [-8.06, -2.20]
3.1 Paediatric patients - severely affected patients	1	64	Mean Difference (IV, Fixed, 95% CI)	-2.60 [-9.68, 4.48]
3.2 Paediatric patients - moderately affected patients	4	712	Mean Difference (IV, Fixed, 95% CI)	-5.65 [-8.87, -2.43]
4 Absence from school or work at 12 months	3	316	Mean Difference (IV, Fixed, 95% CI)	-2.33 [-3.42, -1.24]

Comparison 2. Tonsillectomy or adenotonsillectomy versus no surgery in adults

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Episodes of sore throat of any severity (not including as one episode the period post-surgery) at 6 months	2	156	Mean Difference (IV, Random, 95% CI)	-3.61 [-7.92, 0.70]
2 Sore throat days (not including those immediately post-surgery) at 6 months	2	156	Mean Difference (IV, Fixed, 95% CI)	-10.64 [-15.52, -5.76]

	No. of studies	No. of participants	Statistical method	Effect size
3 Absence from school or work at 6 months	1	96	Mean Difference (IV, Fixed, 95% CI)	-3.3 [-7.66, 1.06]

Comparison 3.   Subgroup analysis: by type of surgery (tonsillectomy or adenotonsillectomy versus no surgery) in children

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Episodes of sore throat of any severity (not including as one episode the period post-surgery)	5	795	Mean Difference (IV, Random, 95% CI)	-0.44 [-0.88, 0.00]
1.1 Adenotonsillectomy	3	417	Mean Difference (IV, Random, 95% CI)	-0.40 [-0.88, 0.07]
1.2 Tonsillectomy	1	74	Mean Difference (IV, Random, 95% CI)	0.18 [-0.53, 0.89]
1.3 Mixed	2	304	Mean Difference (IV, Random, 95% CI)	-1.13 [-1.96, -0.31]
2 Episodes of moderate/severe sore throat (not including as one episode the period post-surgery)	4		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.1 Adenotonsillectomy	3	445	Mean Difference (IV, Fixed, 95% CI)	0.84 [0.72, 0.96]
2.2 Tonsillectomy	1	73	Mean Difference (IV, Fixed, 95% CI)	0.85 [0.65, 1.05]
2.3 Mixed	1	146	Mean Difference (IV, Fixed, 95% CI)	-0.59 [-0.93, -0.25]
3 Sore throat days (including those immediately post-surgery)	5	871	Mean Difference (IV, Fixed, 95% CI)	-5.14 [-8.03, -2.24]
3.1 Adenotonsillectomy	3	468	Mean Difference (IV, Fixed, 95% CI)	-4.51 [-8.09, -0.93]
3.2 Tonsillectomy	1	122	Mean Difference (IV, Fixed, 95% CI)	-5.5 [-13.94, 2.94]
3.3 Mixed	2	281	Mean Difference (IV, Fixed, 95% CI)	-6.75 [-12.81, -0.69]

Comparison 4.   Sensitivity analysis: tonsillectomy or adenotonsillectomy versus no surgery in children - removing Lock 2010

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Episodes of sore throat of any severity (including as one episode the period post-surgery) at 12 months	4	564	Mean Difference (IV, Random, 95% CI)	-0.42 [-0.87, 0.02]
1.1 Paediatric patients - severely affected patients	1	73	Mean Difference (IV, Random, 95% CI)	-0.85 [-1.87, 0.17]
1.2 Paediatric patients - moderately affected patients	3	491	Mean Difference (IV, Random, 95% CI)	-0.36 [-0.86, 0.15]
2 Sore throat days (including those immediately post-surgery) at 12 months	4	559	Mean Difference (IV, Fixed, 95% CI)	-4.26 [-7.29, -1.23]

2.1 Paediatric patients - severely affected patients	1	64	Mean Difference (IV, Fixed, 95% CI)	-2.60 [-9.68, 4.48]
2.2 Paediatric patients - moderately affected patients	3	495	Mean Difference (IV, Fixed, 95% CI)	-4.63 [-7.98, -1.28]

Analysis 1.1. Comparison 1 Tonsillectomy or adenotonsillectomy versus no surgery in children, Outcome 1 Episodes of sore throat of any severity (including as one episode the period post-surgery) at 12 months.

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 1 Tonsillectomy or adenotonsillectomy versus no surgery in children

Outcome: 1 Episodes of sore throat of any severity (including as one episode the period post-surgery) at 12 months

Study or subgroup	Tonsil/adenotonsilectomy		No surgery		Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
	N	Mean(SD)[Episodes]	N	Mean(SD)[Episodes]			
1 Paediatric patients - severely affected patients							
Paradise 1984	38	2.2 (1.62)	35	3.1 (2.64)		14.4 %	-0.90 [-1.92, 0.12]
Subtotal (95% CI)	**38**		**35**			**14.4 %**	**-0.90 [-1.92, 0.12]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.74 (P = 0.082)							
2 Paediatric patients - moderately affected patients							
Lock 2010 (1)	119	6 (5.16)	112	7.7 (5.88)		8.9 %	-1.70 [-3.13, -0.27]
Paradise 2002a	93	2.91 (1.39)	54	2.78 (1.66)		26.4 %	0.13 [-0.40, 0.66]
Paradise 2002b	59	2.9 (1.37)	67	3.6 (1.87)		25.1 %	-0.70 [-1.27, -0.13]
van Staaij 2004	111	2.74 (1.76)	107	3.27 (2.42)		25.2 %	-0.53 [-1.09, 0.03]
Subtotal (95% CI)	**382**		**340**			**85.6 %**	**-0.51 [-1.07, 0.05]**
Heterogeneity: Tau² = 0.19; Chi² = 8.36, df = 3 (P = 0.04); I² =64%							
Test for overall effect: Z = 1.79 (P = 0.074)							
Total (95% CI)	**420**		**375**			**100.0 %**	**-0.56 [-1.04, -0.07]**
Heterogeneity: Tau² = 0.16; Chi² = 9.20, df = 4 (P = 0.06); I² =57%							
Test for overall effect: Z = 2.24 (P = 0.025)							
Test for subgroup differences: Chi² = 0.44, df = 1 (P = 0.51), I² =0.0%							

```
        -4    -2    0    2    4
     Favours surgery   Favours no surgery
```

(1) Lock 2010 has high risk of attrition bias due to the low rates of return of diaries. However, inclusion of data from this analysis did not affect effect sizes (see analysis 4.1).

Analysis 1.2. Comparison 1 Tonsillectomy or adenotonsillectomy versus no surgery in children, Outcome 2 Episodes of moderate/severe sore throat (including as one episode the period post-surgery) at 12 months.

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 1 Tonsillectomy or adenotonsillectomy versus no surgery in children

Outcome: 2 Episodes of moderate/severe sore throat (including as one episode the period post-surgery) at 12 months

Study or subgroup	Tonsil/adenotonsilectomy		No surgery		Mean Difference	Weight	Mean Difference
	N	Mean(SD)[Episodes]	N	Mean(SD)[Episodes]	IV,Random,95% CI		IV,Random,95% CI
1 Paediatric patients - severely affected patients							
Paradise 1984	38	1.08 (0.27)	35	1.17 (1.42)		100.0 %	-0.09 [-0.57, 0.39]
Subtotal (95% CI)	**38**		**35**			**100.0 %**	**-0.09 [-0.57, 0.39]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.37 (P = 0.71)							
2 Paediatric patients - moderately affected patients							
Paradise 2002a	93	1.13 (0.41)	54	0.24 (0.51)		55.7 %	0.89 [0.73, 1.05]
Paradise 2002b	59	1.15 (0.42)	67	0.43 (0.67)		38.1 %	0.72 [0.53, 0.91]
van Staaij 2004	111	1.69 (0.91)	107	0.86 (2.34)		6.3 %	0.83 [0.36, 1.30]
Subtotal (95% CI)	**263**		**228**			**100.0 %**	**0.82 [0.70, 0.94]**
Heterogeneity: Tau² = 0.0; Chi² = 1.77, df = 2 (P = 0.41); I² =0.0%							
Test for overall effect: Z = 13.53 (P < 0.00001)							
Test for subgroup differences: Chi² = 13.14, df = 1 (P = 0.00), I² =92%							

```
          -4    -2     0     2     4
        Favours surgery   Favours no surgery
```
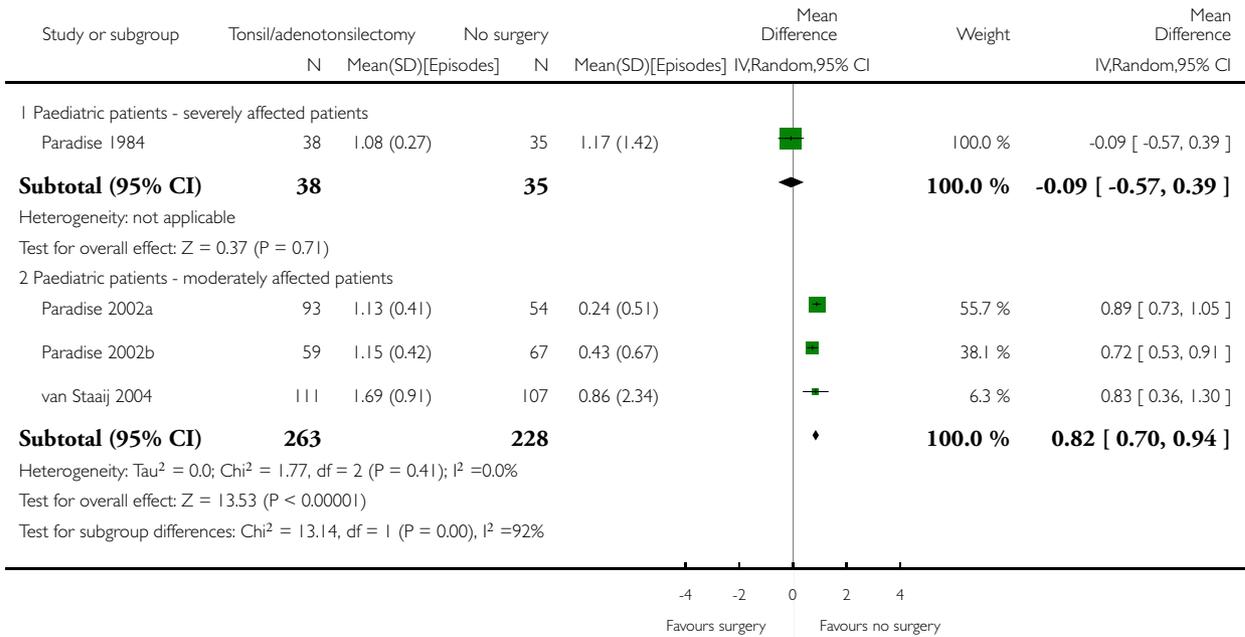
Analysis 1.3.   Comparison 1 Tonsillectomy or adenotonsillectomy versus no surgery in children, Outcome 3 Sore throat days (including those immediately post-surgery) at 12 months.

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 1 Tonsillectomy or adenotonsillectomy versus no surgery in children

Outcome: 3 Sore throat days (including those immediately post-surgery) at 12 months

Study or subgroup	Tonsil/adenotonsillectomy		No surgery		Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
	N	Mean(SD)[Days]	N	Mean(SD)[Days]			
1 Paediatric patients - severely affected patients							
Paradise 1984	31	16.3 (14.3)	33	18.9 (14.6)		17.1 %	-2.60 [-9.68, 4.48]
Subtotal (95% CI)	**31**		**33**			**17.1 %**	**-2.60 [-9.68, 4.48]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.72 (P = 0.47)							
2 Paediatric patients - moderately affected patients							
Lock 2010 (1)	110	31 (29.25)	107	49.1 (54.62)		6.3 %	-18.10 [-29.80, -6.40]
Paradise 2002a	95	19.5 (14.5)	54	25 (21)		21.5 %	-5.50 [-11.81, 0.81]
Paradise 2002b	60	23 (20)	68	24 (17)		20.5 %	-1.00 [-7.48, 5.48]
van Staaij 2004	111	12.68 (9.5)	107	18.92 (24.59)		34.6 %	-6.24 [-11.22, -1.26]
Subtotal (95% CI)	**376**		**336**			**82.9 %**	**-5.65 [-8.87, -2.43]**
Heterogeneity: Chi² = 6.38, df = 3 (P = 0.09); I² =53%							
Test for overall effect: Z = 3.44 (P = 0.00058)							
Total (95% CI)	**407**		**369**			**100.0 %**	**-5.13 [-8.06, -2.20]**
Heterogeneity: Chi² = 6.97, df = 4 (P = 0.14); I² =43%							
Test for overall effect: Z = 3.43 (P = 0.00061)							
Test for subgroup differences: Chi² = 0.59, df = 1 (P = 0.44), I² =0.0%							

```
          -20   -10    0    10    20
        Favours surgery    Favours no surgery
```

(1) Lock 2010 has high risk of attrition bias due to the low rates of return of diaries. Sensitivity analysis (analysis 4.3) showed that mean difference was reduced by about 1 day, and I2=0 (no statistical heterogeneity).

Analysis 1.4. Comparison 1 Tonsillectomy or adenotonsillectomy versus no surgery in children, Outcome 4 Absence from school or work at 12 months.

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 1 Tonsillectomy or adenotonsillectomy versus no surgery in children

Outcome: 4 Absence from school or work at 12 months

Study or subgroup	Tonsil/adenotonsilectomy		No surgery		Mean Difference	Weight	Mean Difference
	N	Mean(SD)[Absences]	N	Mean(SD)[Absences]	IV,Fixed,95% CI		IV,Fixed,95% CI
Paradise 1984	29	3.5 (4.2)	41	6.3 (6.7)		18.2 %	-2.80 [-5.36, -0.24]
Paradise 2002a	86	3.6 (3.8)	50	5.3 (4.7)		50.9 %	-1.70 [-3.23, -0.17]
Paradise 2002b	52	3.5 (4.2)	58	6.6 (6.2)		30.9 %	-3.10 [-5.06, -1.14]
Total (95% CI)	**167**		**149**			**100.0 %**	**-2.33 [-3.42, -1.24]**

Heterogeneity: Chi² = 1.37, df = 2 (P = 0.50); I² =0.0%

Test for overall effect: Z = 4.19 (P = 0.000028)

Test for subgroup differences: Not applicable

```
         -10   -5    0    5    10
        Favours surgery   Favours no surgery
```

Analysis 2.1. Comparison 2 Tonsillectomy or adenotonsillectomy versus no surgery in adults, Outcome 1 Episodes of sore throat of any severity (not including as one episode the period post-surgery) at 6 months.

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 2 Tonsillectomy or adenotonsillectomy versus no surgery in adults

Outcome: 1 Episodes of sore throat of any severity (not including as one episode the period post-surgery) at 6 months

Study or subgroup	Tonsil/adenotonsilectomy		No surgery		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	IV,Random,95% CI		IV,Random,95% CI
Alho 2007	36	0.6 (0.9)	34	2.1 (2.3)		52.0 %	-1.50 [-2.33, -0.67]
Koskenkorva 2013	46	1.5 (2.4)	40	7.4 (5.8)		48.0 %	-5.90 [-7.83, -3.97]
Total (95% CI)	**82**		**74**			**100.0 %**	**-3.61 [-7.92, 0.70]**

Heterogeneity: Tau² = 9.11; Chi² = 16.92, df = 1 (P = 0.00004); I² =94%

Test for overall effect: Z = 1.64 (P = 0.10)

Test for subgroup differences: Not applicable

```
         -10   -5    0    5    10
        Favours surgery   Favours no surgery
```
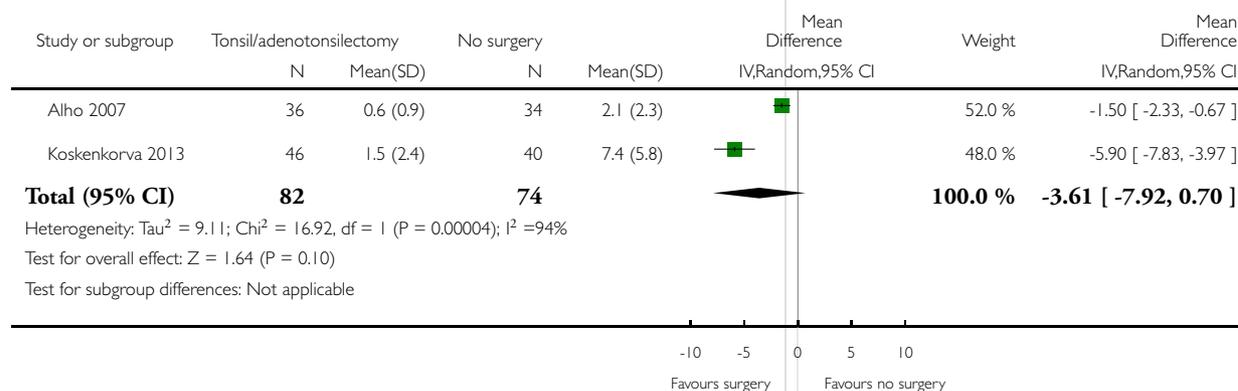
Analysis 2.2.   Comparison 2 Tonsillectomy or adenotonsillectomy versus no surgery in adults, Outcome 2 Sore throat days (not including those immediately post-surgery) at 6 months.

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 2 Tonsillectomy or adenotonsillectomy versus no surgery in adults

Outcome: 2 Sore throat days (not including those immediately post-surgery) at 6 months

Study or subgroup	Tonsil/adenotonsilectomy		No surgery		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	IV,Fixed,95% CI		IV,Fixed,95% CI
Alho 2007	36	3.2 (5.3)	34	12.1 (14.1)		93.4 %	-8.90 [-13.95, -3.85]
Koskenkorva 2013 (1)	46	5.1 (8.7)	40	40.2 (60.5)		6.6 %	-35.10 [-54.02, -16.18]
Total (95% CI)	**82**		**74**			**100.0 %**	**-10.64 [-15.52, -5.76]**

Heterogeneity: Chi² = 6.88, df = 1 (P = 0.01); I² =85%

Test for overall effect: Z = 4.28 (P = 0.000019)

Test for subgroup differences: Not applicable

-20   -10   0   10   20

Favours surgery   Favours no surgery
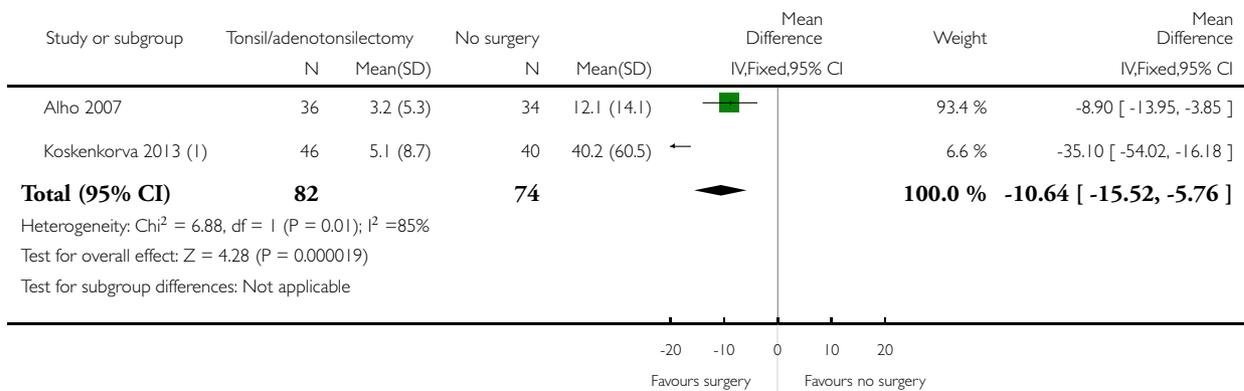
(1) There was a potential for skewed data as standard deviations larger than means were observed.

Analysis 2.3. Comparison 2 Tonsillectomy or adenotonsillectomy versus no surgery in adults, Outcome 3 Absence from school or work at 6 months.

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 2 Tonsillectomy or adenotonsillectomy versus no surgery in adults

Outcome: 3 Absence from school or work at 6 months

Study or subgroup	Tonsil/adenotonsillectomy		No surgery		Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
	N	Mean(SD)	N	Mean(SD)			
Koskenkorva 2013	46	3.3 (10)	50	6.6 (11.8)		100.0 %	-3.30 [-7.66, 1.06]
Total (95% CI)	**46**		**50**			**100.0 %**	**-3.30 [-7.66, 1.06]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.48 (P = 0.14)							
Test for subgroup differences: Not applicable							

```
        -10    -5    0    5    10
        Favours surgery    Favours no surgery
```

Analysis 3.1.   Comparison 3 Subgroup analysis: by type of surgery (tonsillectomy or adenotonsillectomy versus no surgery) in children, Outcome 1 Episodes of sore throat of any severity (not including as one episode the period post-surgery).

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 3 Subgroup analysis: by type of surgery (tonsillectomy or adenotonsillectomy versus no surgery) in children

Outcome: 1 Episodes of sore throat of any severity (not including as one episode the period post-surgery)

Study or subgroup	Surgery N	Mean(SD)	No surgery N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 Adenotonsillectomy							
Paradise 2002a	46	2.96 (1.39)	27	2.78 (1.66)		17.5 %	0.18 [-0.56, 0.92]
Paradise 2002b	59	2.9 (1.37)	67	3.6 (1.87)		22.1 %	-0.70 [-1.27, -0.13]
van Staaij 2004	111	2.74 (1.76)	107	3.27 (2.42)		22.2 %	-0.53 [-1.09, 0.03]
Subtotal (95% CI)	**216**		**201**			**61.8 %**	**-0.40 [-0.88, 0.07]**
Heterogeneity: Tau^2 = 0.08; Chi^2 = 3.57, df = 2 (P = 0.17); I^2 =44%							
Test for overall effect: Z = 1.66 (P = 0.098)							
2 Tonsillectomy							
Paradise 2002a	47	2.96 (1.1449)	27	2.78 (1.66)		18.4 %	0.18 [-0.53, 0.89]
Subtotal (95% CI)	**47**		**27**			**18.4 %**	**0.18 [-0.53, 0.89]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.50 (P = 0.62)							
3 Mixed							
Lock 2010	119	6 (5.16)	112	7.7 (5.88)		7.5 %	-1.70 [-3.13, -0.27]
Paradise 1984	38	2.24 (1.62)	35	3.09 (2.64)		12.3 %	-0.85 [-1.87, 0.17]
Subtotal (95% CI)	**157**		**147**			**19.8 %**	**-1.13 [-1.96, -0.31]**
Heterogeneity: Tau^2 = 0.0; Chi^2 = 0.90, df = 1 (P = 0.34); I^2 =0.0%							
Test for overall effect: Z = 2.69 (P = 0.0072)							
Total (95% CI)	**420**		**375**			**100.0 %**	**-0.44 [-0.88, 0.00]**
Heterogeneity: Tau^2 = 0.15; Chi^2 = 10.12, df = 5 (P = 0.07); I^2 =51%							
Test for overall effect: Z = 1.95 (P = 0.051)							
Test for subgroup differences: Chi^2 = 5.61, df = 2 (P = 0.06), I^2 =64%							

-4   -2   0   2   4

Favours surgery   Favours no surgery

Analysis 3.2. Comparison 3 Subgroup analysis: by type of surgery (tonsillectomy or adenotonsillectomy versus no surgery) in children, Outcome 2 Episodes of moderate/severe sore throat (not including as one episode the period post-surgery).

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 3 Subgroup analysis: by type of surgery (tonsillectomy or adenotonsillectomy versus no surgery) in children

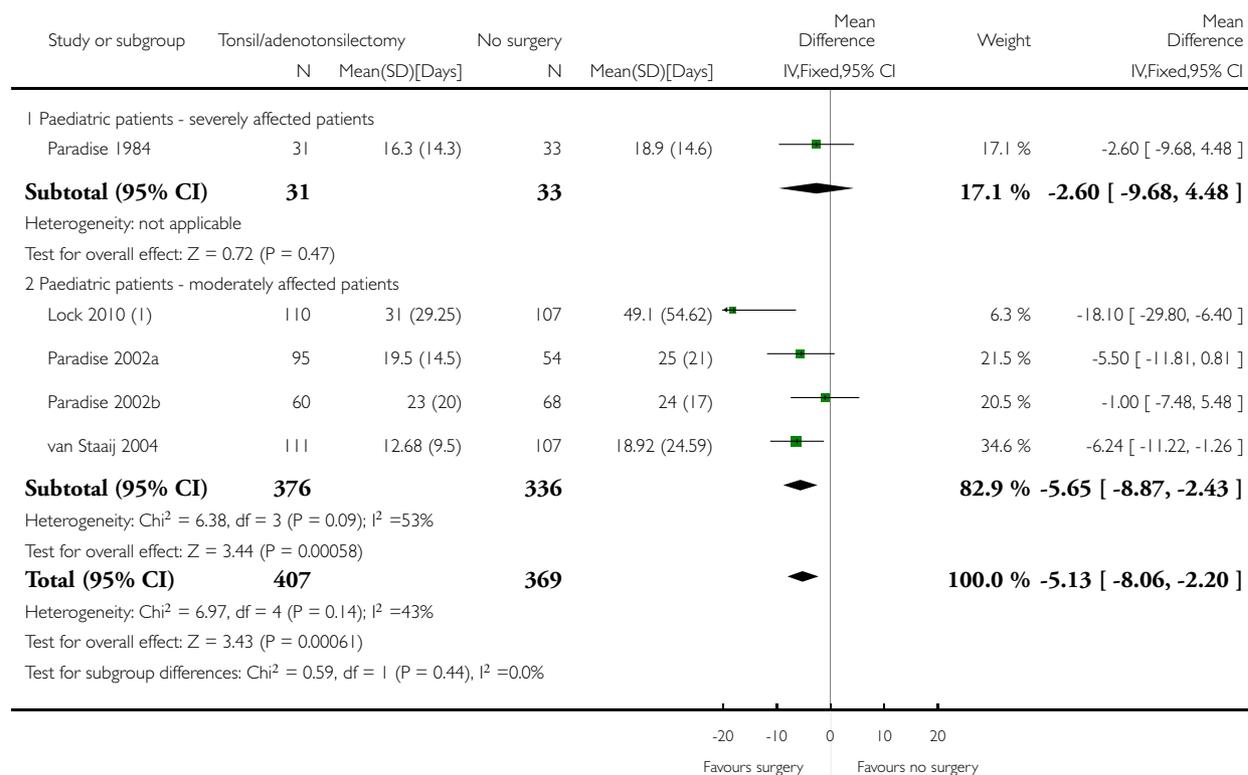Outcome: 2 Episodes of moderate/severe sore throat (not including as one episode the period post-surgery)

Study or subgroup	Surgery N	Mean(SD)	No surgery N	Mean(SD)	Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
1 Adenotonsillectomy							
Paradise 2002a	47	1.17 (0.3406)	54	0.24 (0.51)		53.3 %	0.93 [0.76, 1.10]
Paradise 2002b	59	1.15 (0.42)	67	0.43 (0.67)		40.1 %	0.72 [0.53, 0.91]
van Staaij 2004	111	1.69 (0.91)	107	0.86 (2.34)		6.6 %	0.83 [0.36, 1.30]
Subtotal (95% CI)	**217**		**228**			**100.0 %**	**0.84 [0.72, 0.96]**
Heterogeneity: Chi² = 2.60, df = 2 (P = 0.27); I² =23%							
Test for overall effect: Z = 13.47 (P < 0.00001)							
2 Tonsillectomy							
Paradise 2002a (1)	46	1.09 (0.2357)	27	0.24 (0.51)		100.0 %	0.85 [0.65, 1.05]
Subtotal (95% CI)	**46**		**27**			**100.0 %**	**0.85 [0.65, 1.05]**
Heterogeneity: not applicable							
Test for overall effect: Z = 8.16 (P < 0.00001)							
3 Mixed							
Paradise 1984	38	0.08 (0.27)	35	1.17 (1.42)		50.0 %	-1.09 [-1.57, -0.61]
Paradise 1984	38	1.08 (0.27)	35	1.17 (1.42)		50.0 %	-0.09 [-0.57, 0.39]
Subtotal (95% CI)	**76**		**70**			**100.0 %**	**-0.59 [-0.93, -0.25]**
Heterogeneity: Chi² = 8.40, df = 1 (P = 0.004); I² =88%							
Test for overall effect: Z = 3.42 (P = 0.00063)							
Test for subgroup differences: Chi² = 62.87, df = 2 (P = 0.00), I² =97%							

```
        -4    -2    0    2    4
     Favours surgery    Favours no surgery
```

(1) The

Analysis 3.3.   Comparison 3 Subgroup analysis: by type of surgery (tonsillectomy or adenotonsillectomy versus no surgery) in children, Outcome 3 Sore throat days (including those immediately post-surgery).
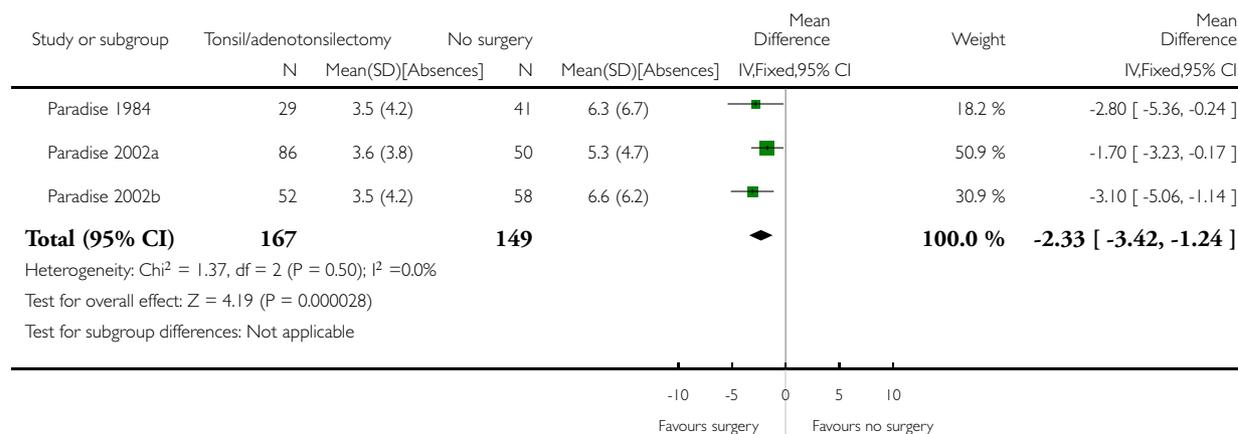
Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 3 Subgroup analysis: by type of surgery (tonsillectomy or adenotonsillectomy versus no surgery) in children

Outcome: 3 Sore throat days (including those immediately post-surgery)

Study or subgroup	Surgery		No surgery		Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
	N	Mean(SD)	N	Mean(SD)			
1 Adenotonsillectomy							
Paradise 2002a	95	19.5 (14.5)	27	25 (21)		11.8 %	-5.50 [-13.94, 2.94]
Paradise 2002b	60	23 (20)	68	24 (17)		20.0 %	-1.00 [-7.48, 5.48]
van Staaij 2004	111	12.68 (9.5)	107	18.92 (24.59)		33.7 %	-6.24 [-11.22, -1.26]
Subtotal (95% CI)	**266**		**202**			**65.4 %**	**-4.51 [-8.09, -0.93]**
Heterogeneity: Chi² = 1.64, df = 2 (P = 0.44); I² =0.0%							
Test for overall effect: Z = 2.47 (P = 0.014)							
2 Tonsillectomy							
Paradise 2002a	95	19.5 (14.5)	27	25 (21)		11.8 %	-5.50 [-13.94, 2.94]
Subtotal (95% CI)	**95**		**27**			**11.8 %**	**-5.50 [-13.94, 2.94]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.28 (P = 0.20)							
3 Mixed							
Lock 2010	110	31 (29.25)	107	49.1 (54.62)		6.1 %	-18.10 [-29.80, -6.40]
Paradise 1984	31	16.3 (14.3)	33	18.9 (14.6)		16.7 %	-2.60 [-9.68, 4.48]
Subtotal (95% CI)	**141**		**140**			**22.8 %**	**-6.75 [-12.81, -0.69]**
Heterogeneity: Chi² = 4.93, df = 1 (P = 0.03); I² =80%							
Test for overall effect: Z = 2.18 (P = 0.029)							
Total (95% CI)	**502**		**369**			**100.0 %**	**-5.14 [-8.03, -2.24]**
Heterogeneity: Chi² = 6.98, df = 5 (P = 0.22); I² =28%							
Test for overall effect: Z = 3.48 (P = 0.00050)							
Test for subgroup differences: Chi² = 0.40, df = 2 (P = 0.82), I² =0.0%							

```
        -20   -10    0    10    20
     Favours surgery    Favours no surgery
```

Analysis 4.1. Comparison 4 Sensitivity analysis: tonsillectomy or adenotonsillectomy versus no surgery in children - removing Lock 2010, Outcome 1 Episodes of sore throat of any severity (including as one episode the period post-surgery) at 12 months.

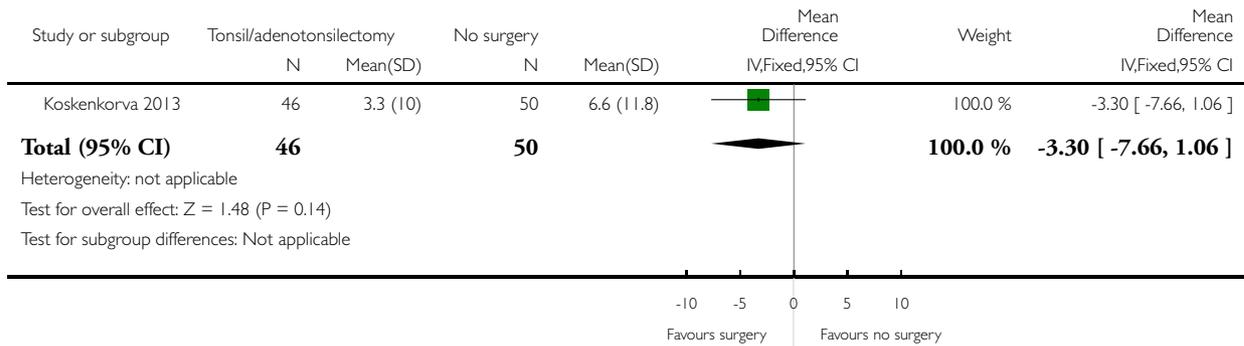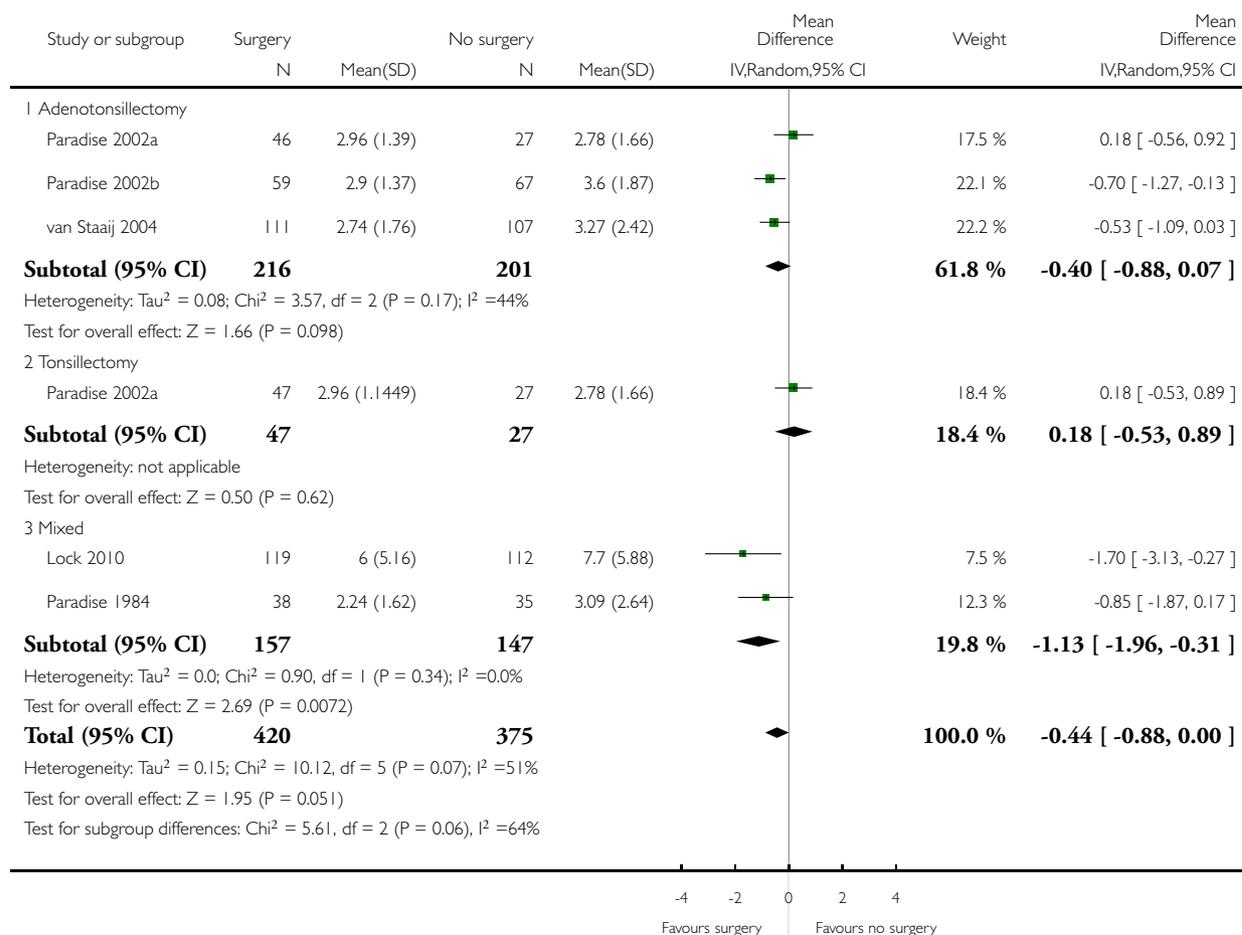Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis

Comparison: 4 Sensitivity analysis: tonsillectomy or adenotonsillectomy versus no surgery in children - removing Lock 2010

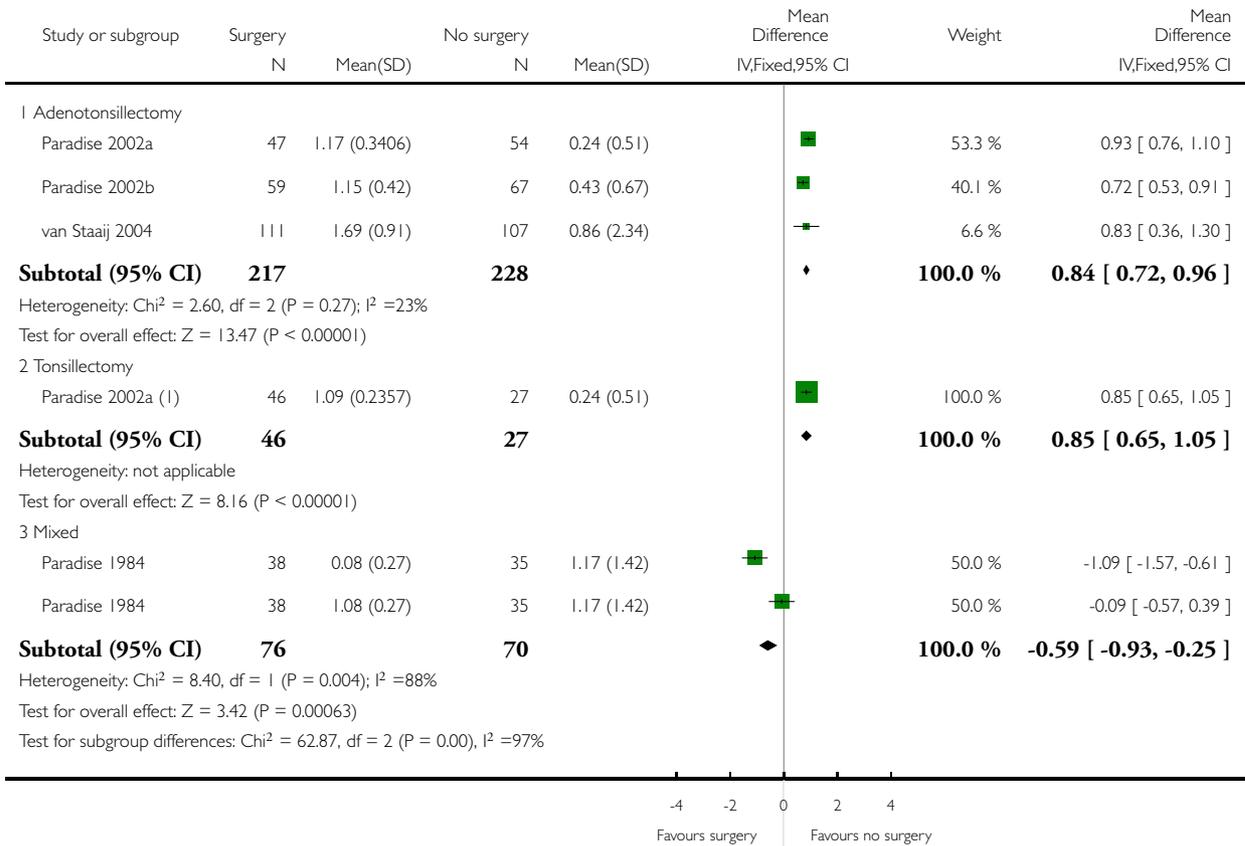Outcome: 1 Episodes of sore throat of any severity (including as one episode the period post-surgery) at 12 months

Study or subgroup	Tonsil/adenotonsilectomy N	Mean(SD)	No surgery N	Mean(SD)	Mean Difference IV,Random,95% CI	Weight	Mean Difference IV,Random,95% CI
1 Paediatric patients - severely affected patients							
Paradise 1984	38	2.24 (1.62)	35	3.09 (2.64)		13.9 %	-0.85 [-1.87, 0.17]
Subtotal (95% CI)	**38**		**35**			**13.9 %**	**-0.85 [-1.87, 0.17]**
Heterogeneity: not applicable							
Test for overall effect: Z = 1.64 (P = 0.10)							
2 Paediatric patients - moderately affected patients							
Paradise 2002a	93	2.91 (1.39)	54	2.78 (1.66)		30.0 %	0.13 [-0.40, 0.66]
Paradise 2002b	59	2.9 (1.37)	67	3.6 (1.87)		28.0 %	-0.70 [-1.27, -0.13]
van Staaij 2004	111	2.74 (1.76)	107	3.27 (2.42)		28.2 %	-0.53 [-1.09, 0.03]
Subtotal (95% CI)	**263**		**228**			**86.1 %**	**-0.36 [-0.86, 0.15]**
Heterogeneity: Tau² = 0.12; Chi² = 5.06, df = 2 (P = 0.08); I² =60%							
Test for overall effect: Z = 1.38 (P = 0.17)							
Total (95% CI)	**301**		**263**			**100.0 %**	**-0.42 [-0.87, 0.02]**
Heterogeneity: Tau² = 0.10; Chi² = 5.94, df = 3 (P = 0.11); I² =49%							
Test for overall effect: Z = 1.88 (P = 0.060)							
Test for subgroup differences: Chi² = 0.73, df = 1 (P = 0.39), I² =0.0%							

-10   -5   0   5   10

Favours surgery   Favours no surgery

Review: Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis
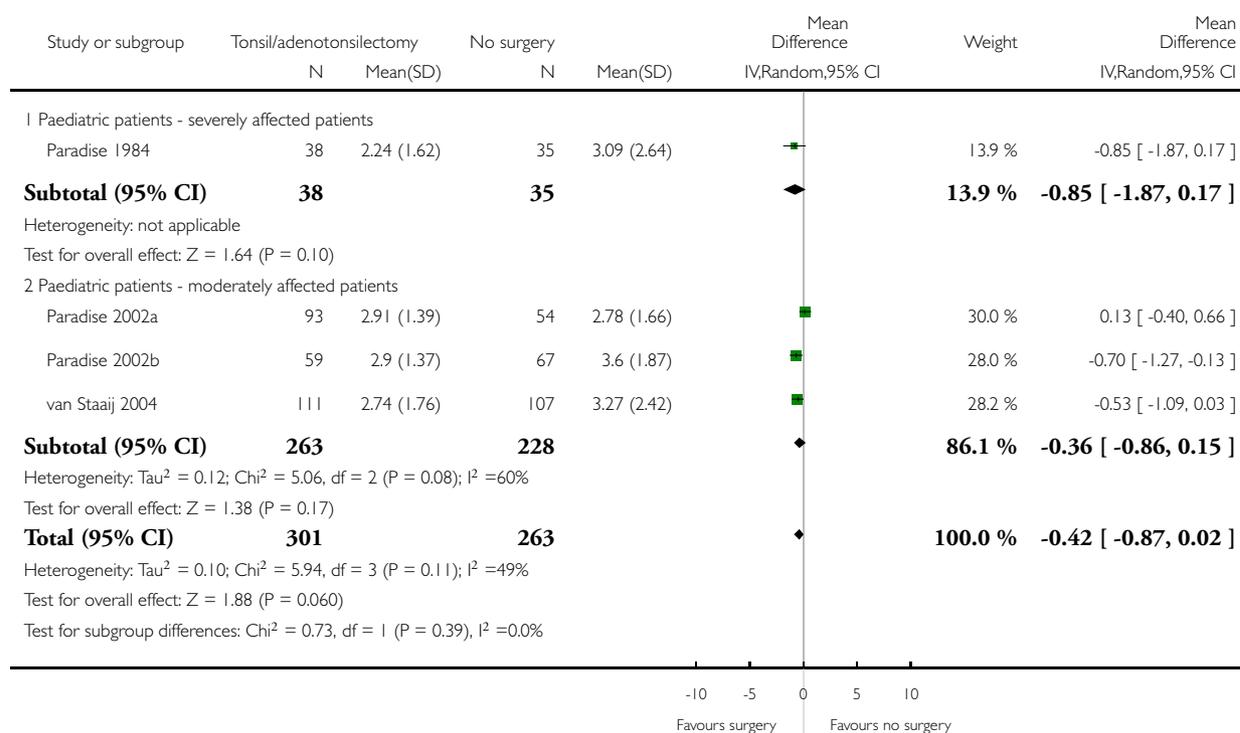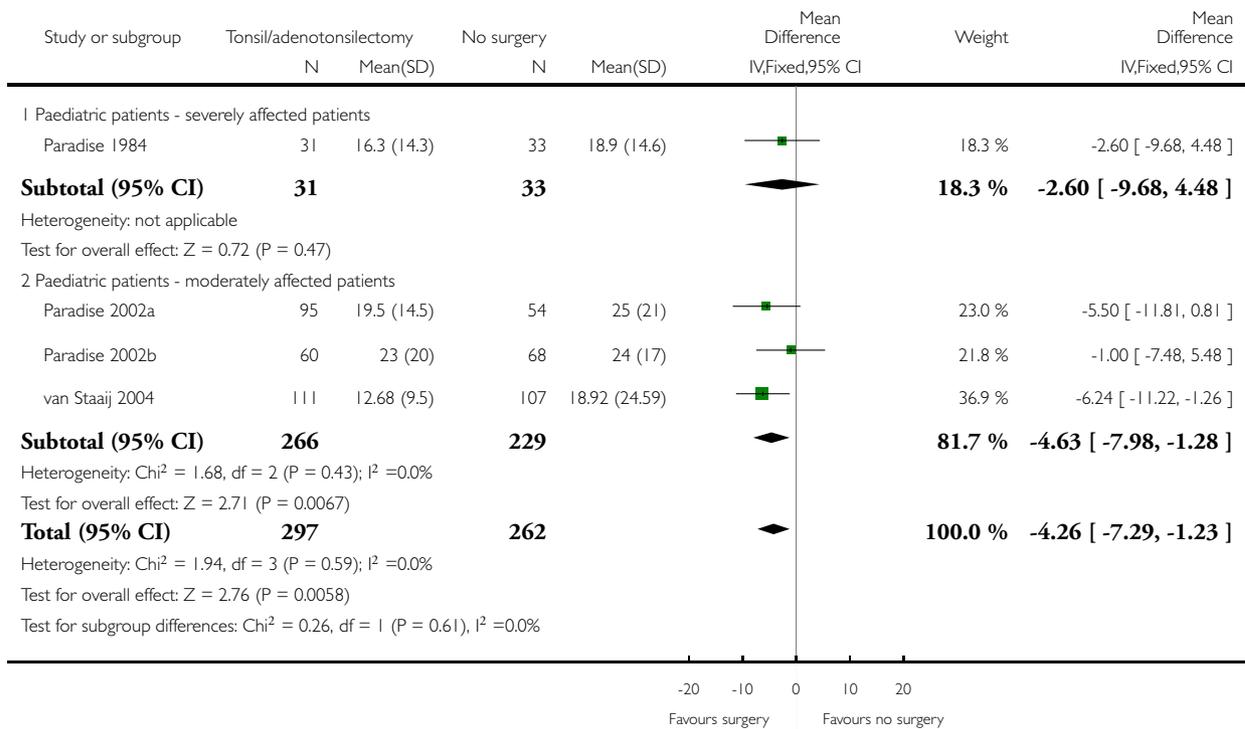
Comparison: 4 Sensitivity analysis: tonsillectomy or adenotonsillectomy versus no surgery in children - removing Lock 2010

Outcome: 2 Sore throat days (including those immediately post-surgery) at 12 months

Study or subgroup	Tonsil/adenotonsilectomy		No surgery		Mean Difference IV,Fixed,95% CI	Weight	Mean Difference IV,Fixed,95% CI
	N	Mean(SD)	N	Mean(SD)			
1 Paediatric patients - severely affected patients							
Paradise 1984	31	16.3 (14.3)	33	18.9 (14.6)		18.3 %	-2.60 [-9.68, 4.48]
Subtotal (95% CI)	**31**		**33**			**18.3 %**	**-2.60 [-9.68, 4.48]**
Heterogeneity: not applicable							
Test for overall effect: Z = 0.72 (P = 0.47)							
2 Paediatric patients - moderately affected patients							
Paradise 2002a	95	19.5 (14.5)	54	25 (21)		23.0 %	-5.50 [-11.81, 0.81]
Paradise 2002b	60	23 (20)	68	24 (17)		21.8 %	-1.00 [-7.48, 5.48]
van Staaij 2004	111	12.68 (9.5)	107	18.92 (24.59)		36.9 %	-6.24 [-11.22, -1.26]
Subtotal (95% CI)	**266**		**229**			**81.7 %**	**-4.63 [-7.98, -1.28]**
Heterogeneity: Chi² = 1.68, df = 2 (P = 0.43); I² =0.0%							
Test for overall effect: Z = 2.71 (P = 0.0067)							
Total (95% CI)	**297**		**262**			**100.0 %**	**-4.26 [-7.29, -1.23]**
Heterogeneity: Chi² = 1.94, df = 3 (P = 0.59); I² =0.0%							
Test for overall effect: Z = 2.76 (P = 0.0058)							
Test for subgroup differences: Chi² = 0.26, df = 1 (P = 0.61), I² =0.0%							

-20 -10 0 10 20

Favours surgery   Favours no surgery

ADDITIONAL TABLES

Table 1. Complications from tonsillectomy/adenotonsillectomy

Study	Type of complication	Number of patients with complications
Paediatric patients		
Lock 2010	Not reported (participants asked to record in a diary whether a list of symptoms - sore ear, difficulty swallowing, nausea and vomiting, aches and pain etc. - was present)	
Paradise 2002a and Paradise 2002b	Total complication rates reported	16/203 (7.9%) reported

	Haemorrhage, intraoperative	3/203 (1.5%)
	Haemorrhage, 2 days to 2 weeks post-surgery	7/203 (3.4%)
	Anaesthetic induction trismus and 'possible' incipient malignant hyperthermia	1/203 (0.5%)
	Postoperative throat pain (days)	Mean 6.3 (range 0 to 21 days)
Paradise 1984	Haemorrhage - "before discharge"	2/95 (2%)
	Haemorrhage - "after discharge"	2/95 (2%)
van Staaij 2004	Total complication rates reported	12/145 (6%)
	Primary haemorrhage	7/145 (4.8%)
	Postoperative nausea	5/145 (3.4%)
Adults		
Alho 2007	"Mild" secondary bleeding	2/36 (6%)
	Postoperative throat pain (days)	Mean 13 (SD 4)
Koskenkorva 2013	Readmission into hospital due to secondary bleeding	2/46 (4.3%)
	Readmission into hospital due to severe pain	1/46 (2.1%)

SD: standard deviation

APPENDICES

Appendix 1. Revised search strategies used since January 2011

CENTRAL	PubMed	EMBASE (Ovid)
#1 MeSH descriptor Tonsillectomy explode all trees #2 MeSH descriptor Palatine Tonsil explode all trees with qualifier: SU #3 tonsillectom*:ti OR tonsilectom*:ti OR adenotonsillectom*:ti OR adeno-tonsillectom*:ti #4 (tonsil*:ti OR adenotonsil*:ti) AND (surg*:ti OR laser*:ti OR extract*:ti OR resect*:ti OR excis*:ti OR operat*:ti OR dissect*:ti OR remov*:ti OR coblat:ti* OR ablat*:ti) #5 (#1 OR #2 OR #3 OR #4)	#1 "Tonsillectomy" [Mesh] OR tonsillectom* [ti] OR tonsilectom* [ti] OR adenotonsillectom* [ti] OR adeno-tonsillectom* [ti] #2 "Palatine Tonsil/surgery"[Mesh] #3 (Tonsil* [ti] OR adenotonsil* [ti]) AND (SURG* [ti] OR OPERAT* [ti] OR EXCIS* [ti] OR EXTRACT* [ti] OR REMOV* [ti] OR DISSECT* [ti] OR ABLAT* [ti] OR COBLAT* [ti] OR LASER* [ti]) #4 #1 OR #2 OR #3	1 exp *tonsillectomy/ 2 exp tonsil/su [Surgery] 3 (tonsillectom* or tonsilectom* or adenotonsillectom* or adeno-tonsillectom*).ti. 4 ((Tonsil* or adenotonsil*) and (SURG* or OPERAT* or EXCIS* or EXTRACT* or REMOV* or DISSECT* or ABLAT* or COBLAT* or LASER*)).ti. 5 1 or 2 or 3 or 4

CINAHL (EBSCO)	Web of Science (Web of Knowledge)	ICTRP
S1 (MH "Tonsillectomy") S2 TI tonsillectom* OR tonsilectom* OR adenotonsillectom* OR adeno-tonsillectom* S3 TI (tonsil* OR adenotonsil*) AND (surg* OR laser* OR extract* OR resect* OR excis* OR operat* OR dissect* OR remov* OR coblat* OR ablat*) S4 S1 or S2 or S3	#1 TI=(tonsillectom* or tonsilectom* or adenotonsillectom* or adeno-tonsillectom*) #2 TI=((Tonsil* or adenotonsil*) and (SURG* or OPERAT* or EXCIS* or EXTRACT* or REMOV* or DISSECT* or ABLAT* or COBLAT* or LASER*)) #3 #1 OR #2	tonsillectom* OR tonsilectom* OR adenotonsillectom* OR adeno-tonsillectom*

Appendix 2. Search strategies used prior to January 2011

CENTRAL	MEDLINE	EMBASE
#1 TONSILLECTOMY single term (MeSH) #2 TONSIL [su] single term (MeSH) #3 TONSILLECTOM* OR TONSILECTOM* #4 ADENOTONSILLECTOM* OR ADENOTONSILECTOM* #5 #1 OR #2 OR #3 OR #4 #6 TONSILLITIS explode all trees (MeSH)	1. TONSILLECTOMY.DE. 2. TONSIL-SU.DE. 3. (TONSILLECTOM$3 OR TONSILECTOM3).TI,AB. 4. (ADENOTONSILLECTOM$3 OR ADENOTONSILECTOM$3).TI,AB. 5. 1 OR 2 OR 3 OR 4 6. TONSILLITIS#.DE. 7. TONSIL.DE. 8. (TONSIL$5 OR ADENOTONSIL$5)	1. TONSILLECTOMY#.DE. 2. TONSIL-DISEASE-SU#.DE. 3. (TONSILLECTOM$3 OR TONSILECTOM$3).TI,AB. 4. (ADENOTONSILLECTOM$3 OR ADENOTONSILECTOM$3).TI,AB. 5. 1 OR 2 OR 3 OR 4 6. TONSILLITIS#.DE. 7. TONSIL.DE. 8. (TONSIL$5 OR ADENOTONSIL$5)

#7 TONSIL single term (MeSH) #8 TONSIL* OR ADENOTONSIL* #9 #6 OR #7 OR #8 #10 SURGERY single term (MeSH) #11 SURGICAL PROCEDURES ELECTIVE single term (MeSH) #12 OTORHINOLARYNGOLOGIC SURGICAL PROCEDURES single term (MeSH) #13 OTOLARYNGOLOGY single term (MeSH) #14 SURG* OR OPERAT* OR EXCIS* OR EXTRACT* OR REMOV* OR DISSECT* OR ABLAT* OR COBLAT* OR LASER* #15 #10 OR #11 OR #12 OR #13 OR #14 #16 #9 AND #15 #17 #5 OR #16	.TI,AB. 9. 6 OR 7 OR 8 10. SURGERY.DE. 11. SURGICAL-PROCEDURES-ELECTIVE.DE. 12. OTORHINOLARYNGOLOGIC-SURGICAL-PROCEDURES.DE. 13. OTOLARYNGOLOGY.DE. 14. (SURG$6 OR OPERAT$4 OR EXCIS$4 OR EXTRACT$4 OR REMOV$3 OR DISSECT$3 OR ABLAT$3 OR COBLAT$3 OR LASER$3).TI,AB. 15. 10 OR 11 OR 12 OR 13 OR 14 16. 9 AND 15 17. 5 OR 16	.TI,AB. 9. 6 OR 7 OR 8 10. SURGERY.DE. 11. SURGICAL-TECHNIQUE#.DE. 12. THROAT-SURGERY.DE. 13. EAR-NOSE-THROAT-SURGERY.DE. 14. (SURG$6 OR OPERAT$4 OR EXCIS$4 OR EXTRACT$4 OR REMOV$3 OR DISSECT$3 OR ABLAT$3 OR COBLAT$3 OR LASER$3).TI,AB. 15. 10 OR 11 OR 12 OR 13 OR 14 16. 9 AND 15 17. 5 OR 16

Appendix 3. Individual results of included studies

Study 1: Paradise 1984

The results of the Paradise 1984 trial are based on surgical intervention in 43 severely affected children compared to 48 control subjects. Of those having surgery, 27 underwent tonsillectomy alone and 16 adenotonsillectomy; the results for the two procedures are pooled and inseparable.

A: Episodes of sore throat of any severity

For the mean number of episodes of sore throats of any severity within 12 months the results are:

Adeno-/tonsillectomy group (n = 38)	1.2 episodes (SD 1.62)
Control group (n = 35)	3.1 episodes (SD 2.64)

However, we had to assume that the data had not included the episode post-surgery and, therefore, the number of episodes of sore throat in the surgery group should be 2.2 episodes (SD 1.62).

B: Episodes of moderate/severe sore throat

The numbers of moderate or severe sore throats in the first year of the trial were as follows:

Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Review)   **67**

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Adeno-/tonsillectomy group (n = 38)	0.1 episodes (SD 0.27)
Control group (n = 35)	1.2 episodes (SD 1.42)

We must **assume** that the data for the surgical groups cannot include the immediate postoperative period as one episode of sore throat (because if they did, the mean in the surgical group would have to be > 1). Therefore, the number of episodes of moderate to severe sore throat should be 1.1 (SD 0.27) in the surgery group.

C: Sore throat days (including those immediately following surgery)

Data on the number of sore throat days in the surgical and control groups are only available for 31 and 33 children respectively (72% and 69% of those enrolled) (see Paradise 1984, Table 5). Here sore throat days immediately after surgery are included and "the number of days for each subject for each follow-up year was standardised on the basis of 365 days". Data on the mean number of sore throat days are as follows:

Adeno-/tonsillectomy group (n = 31)	16.3 days (SD 14.3)
Control group (n = 33)	18.9 days (SD 14.6)

The number of days with sore throat postoperatively varies considerably. In the Paradise 1984 trial a mean figure of 4.9 days is reported. In the later studies there is a mean of 6.3 days with a wide range from 0 to 21 days (Paradise 2002a; Paradise 2002b). Furthermore, it is interesting to note that despite the fact that the trial included only children who were severely affected by throat infections, following enrolment in the trial many of those in the control (non-operated) group had few episodes of infection and these few were usually mild. Of the 48 children in the original control group, seven (15%) had elected to have surgery before the end of the first year and were excluded from analysis. Another six children were excluded because of loss to follow-up (n = 4) or not completing the whole 12-month period (n = 2). Of the 35 remaining children, however, 26 (74% of that group; 54% of the whole control group) had either a single episode of moderate or severe sore throat or none at all.

Adverse events associated with surgery

Of the 95 children treated with surgery (children of the randomised and non-randomised studies combined), 13 (14%) had surgery-related complications and six of them (46% of the children that suffered from surgery-related complications; 6% of the children treated with surgery) required one or more extra days in the hospital. The primary and secondary haemorrhage rates were both 2%, of which none required transfusion.

Study 2: Paradise 2002a

In the Paradise 2002a; Paradise 2002b trials, the authors present their results in a slightly different way. Data are presented for the number of episodes in the first 12 months (Paradise 2002a; Paradise 2002b, Table 2). However, the mean number of episodes and the range is given, along with the 95% confidence interval for the mean. For the purposes of this review, the standard deviation of the mean has had to be imputed, by dividing the confidence interval by four and multiplying by the $\sqrt{(N-1)}$, where N is the number of subjects in the relevant group.

The results of the Paradise 2002a trial are based on surgical intervention in 117 children less severely affected than those included in Paradise 1984 (tonsillectomy: 58 children; adenotonsillectomy: 59 children) compared to 60 control subjects. Results for the two procedures are presented separately. The results are presented as the mean number of episodes and sore throat days with accompanying standard deviations.

A: Episodes of sore throat of any severity

For the mean number of episodes of "all [degrees of sore throat] combined" in the first year results are (*SD imputed as described):

Tonsillectomy group (n = 47)	2.0 episodes (95% CI 1.58 to 2.40; SD* 1.39)
Adenotonsillectomy group (n = 46)	1.9 episodes (95% CI 1.51 to 2.34; SD* 1.39)
Combined: tonsillectomy and adenotonsillectomy (n = 93)	1.9 episodes (SD 1.39)
Control group (n = 54)	2.8 episodes (95% CI 2.35 to 3.26; SD* 1.66)

However, we noted that the number of episodes was unlikely to include the episode of pain due to the surgery. Therefore, to include the episode of pain due to surgery, the analysis considered the episodes of sore throat in the tonsillectomy group as 3.0 episodes (SD* 1.39) in the tonsillectomy group and 2.9 episodes (SD* 1.39) in the adenotonsillectomy group.

B: Episodes of moderate/severe sore throat

For the mean number of episodes of moderate or severe sore throat in the first year results are (*SD imputed as described):

Tonsillectomy group (n = 47)	0.2 episodes (95% CI 0.07 to 0.34; SD* 0.46)
Adenotonsillectomy group (n = 46)	0.1 episodes (95% CI 0.02 to 0.23; SD* 0.35)
Combined: tonsillectomy and adenotonsillectomy (n = 93)	0.1 episodes (SD 0.41)
Control group (n = 54)	0.2 episodes (95% CI 0.13 to 0.41; SD* 0.51)

The data for the surgical groups cannot include the immediate postoperative period as one episode of sore throat (see above). The clinical interpretation therefore must be that the 'benefit' of avoiding 0.9 episodes of any sore throat is offset by the episode of sore throat attributable to the surgery, whilst there was no benefit in terms of a reduction in the number of episodes of moderate/severe sore throat after the immediate postoperative period, despite having experienced the postoperative episode of pain that naturally follows the surgery. To include the episode of pain due to surgery in the analysis, the number of episodes of moderate or severe sore throat was considered as 1.2 (SD 0.46) in the tonsillectomy group and 1.1 (SD 0.35) in the adenotonsillectomy group.

C: Sore throat days (including those immediately following surgery)

Data are available on the number of sore throat days in the first 12 months. These are presented as the mean number of sore throat days and standard deviations. As in the Paradise 1984 trial, the data are "limited to subjects with at least 270 days of reportage in a follow-up year", the sore throat days immediately after surgery are included and "the number of days for each subject for each follow-up year was standardised on the basis of 365 days". Data are available for slightly more children:

Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Review) 69

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Tonsillectomy group (n = 48)	20 days (SD 14)
Adenotonsillectomy group (n = 47)	19 days (SD 15)
Combined: tonsillectomy and adenotonsillectomy (n = 95)	20 days (SD 14.5)
Control group (n = 54)	25 days (SD 21)

Adverse events associated with surgery

Of the 203 children who underwent surgery (in the entire study, Paradise 2002a and Paradise 2002b combined) 16 (7.9%) had intra- or postoperative complications. The primary and secondary haemorrhage rates were 1.5% and 3.4%, respectively.

Study 3: Paradise 2002b

The results of the Paradise 2002b trial are based on surgical intervention (adenotonsillectomy) in 73 less severely affected children than those included in Paradise 1984 compared to 78 control subjects. The results are presented in an identical way to Paradise 2002a, as the mean number of episodes and sore throat days with accompanying standard deviations and by intention-to-treat.

A: Episodes of sore throat of any severity

For the mean number of episodes of "all [degrees of sore throat] combined" in the first year results are (* SD imputed as described):

Adenotonsillectomy group (n = 59)	1.9 episodes (95% CI 1.56 to 2.28; SD* 1.37)
Control group (n = 67)	3.6 episodes (95% CI 3.16 to 4.08; SD* 1.87)

As the data were unlikely to account for the episode due to surgery, we added one episode to the surgery group. The mean number of episodes should be 2.9 SD 1.37.

B: Episodes of moderate/severe sore throat

For the mean number of episodes of moderate or severe sore throat in the first year the results are (*SD imputed as described):

Adenotonsillectomy group (n = 59)	0.2 episodes (95% CI 0.07 to 0.29; SD* 0.42)
Control group (n = 67)	0.4 episodes (95% CI 0.29 to 0.62; SD* 0.67)

The data for the surgical groups cannot include the immediate postoperative period as one episode of sore throat (see above). Therefore, for the purpose of data analysis, we added one episode to the surgery group; the mean number of episodes should be 1.2 (SD 0.42).

C: Sore throat days (including those immediately following surgery)

Data are available on the number of sore throat days in the first 12 months. These are presented as the mean number of sore throat days and standard deviations. As in the Paradise 1984 trial, data are "limited to subjects with at least 270 days of reportage in a follow-up year", the sore throat days immediately after surgery are included and "the number of days for each subject for each follow-up year was standardised on the basis of 365 days". Data are available for slightly more children:

Adenotonsillectomy group (n = 60)	23 days (SD 20)
Control group (n = 68)	24 days (SD 17)

The difference between the adenotonsillectomy group and the control group is not statistically significant: -1 day (95% CI -7.48 to 5.48), and the issues of timing and predictability apply again.

Adverse events associated with surgery

See above.

Study 4: van Staaij 2004

The primary outcome measure in this trial was the incidence and duration of fever (temperature of 38.0 °C or higher) for at least one day. Secondary outcome measures included throat infections (fever plus sore throat or pain or difficulty swallowing) and sore throats (sore throat or pain or difficulty swallowing with or without fever). Data relating to outcomes at 12 months, comparable to those in the Paradise studies (Paradise 1984; Paradise 2002a; Paradise 2002b), are not presented in the paper reporting the results of this trial but have been obtained from the trialists.

The results of the van Staaij 2004 trial are based on surgical intervention (adenotonsillectomy) in 151 children with mild to moderate symptoms compared to 149 control (watchful waiting) subjects. The results are presented as the mean number of episodes and sore throat days with accompanying standard deviations and according the group randomised. (Results from the authors).

A: Episodes of sore throat of any severity

The number of episodes of sore throat in the first year of the trial, including as one of these episodes the sore throat immediately following surgery (unlike the Paradise studies), were as follows:

Adenotonsillectomy group (n = 111)	2.7 episodes (SD 1.76)
Control group (n = 107)	3.3 episodes (SD 2.42)

B: Episodes of moderate/severe sore throat

The number of episodes of throat infection in the first year of the trial, were as follows:

Adenotonsillectomy group (n = 111)	0.7 episodes (SD 0.91)
Control group (n = 107)	0.9 episodes (SD 1.12)

In this case, the episode of sore throat following surgery cannot have routinely been included as an episode of 'throat infection' (see above). The difference between the adenotonsillectomy group and the control group in the number of episodes of moderate/severe sore throat is not statistically significant: -0.2 episodes (95% CI -0.64 to 0.30).

To allow a similar comparison with other data, the episode due to surgery should be included. Therefore, the number of episodes of moderate or severe pain should be 1.7 (SD 0.91) in the adenotonsillectomy group.

C: Sore throat days (including those immediately following surgery)

The number of sore throat days (also including sore throat days immediately following surgery) in the first year of the trial was:

Adenotonsillectomy group (n = 111)	12.7 days (SD 9.5)
Control group (n = 107)	18.9 days (SD 24.59)

The difference between the adenotonsillectomy group and the control treatment group is statistically significant: -6.2 days (95% CI -11.22 to -1.26). The issues of timing and predictability apply equally here.

Van Staaij 2004 - year two data

The study has relatively low loss to follow-up overall. However, not all patients completed two years of follow-up as the study had to stop at a predetermined time point (February 2003). Overall, losses to follow-up were relatively low for this study: 88% (133/151) and 83% (124/149) of the adenotonsillectomy and control groups respectively. However, the study was designed to complete at a predetermined time point (February 2003) and, therefore, not all patients could complete up to 2 years of follow-up.

The authors confirmed that whilst there were 300 randomised participants, 24-month data were only *potentially* available for a group of 166 who had been in the trial long enough to complete 24 months follow-up. In fact, data were only available for 80 of these 166 participants. Data on the occurrence of sore throats in the second year (months 13 to 24) were not available.

The *published* data (referring to rates *per year*) record the incidence of symptoms at a median follow-up interval of 22 months. These are:

A: Episodes of sore throat of any severity

Adenotonsillectomy group	2.25 episodes per year	4.5 episodes over 2 years
Control group	2.85 episodes per year	5.7 episodes over 2 years

B: Episodes of moderate/severe sore throat

Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Review) 72

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

| Adenotonsillectomy group | 0.56 episodes per year | 1.12 episodes over 2 years |
| Control group | 0.77 episodes per year | 1.54 episodes over 2 years |

C: Sore throat days (including those immediately following surgery)

| Adenotonsillectomy group | 9.81 days per year | 19.62 days over 2 years |
| Control group | 15.71 days per year | 31.42 days over 2 years |

Each of these *published* comparisons reveal statistically significant differences in:
(A) episodes of any sore throat per year of -0.60 episodes (95% CI -0.90 to -0.30);
(B) episodes of moderate/severe sore throat per year of -0.21 (95% CI -0.36 to -0.06);
(C) sore throat days of -5.91 (95% CI -6.57 to -5.24) in favour of the adenotonsillectomy group.
Again, note that some of these data include sore throats in the immediate postoperative period.

Adverse events associated with surgery
Of the 195 children who underwent surgery (145 in the adenotonsillectomy group and 50 in the watchful waiting group) 12 (6%) had complications related to surgery. The primary haemorrhage rate was 4% (secondary rate not given).

Study 5: Lock 2010
The primary outcome measure in this trial was the reported number of episodes of sore throat. Secondary outcomes included reported number of episodes of sore throat that invoked a GP consultation and number of days with sore throat.
The results of the Lock 2010 trial are based on surgical intervention (tonsillectomy alone or adenotonsillectomy) in 131 children with recurrent acute sore throat compared to 137 control subjects. In the control group, 26% (36/137) elected to undergo surgery during follow-up, whereas 8% (11/131) children allocated to adeno-/tonsillectomy did not receive surgery. The results are presented as the mean number of episodes and sore throat days with accompanying standard deviations and by intention-to-treat.

A: Episodes of sore throat of any severity
The number of episodes of sore throat in the first year of the trial, <u>including</u> as one of these episodes the sore throat immediately following surgery (unlike the Paradise studies), was as follows:

| Adeno-/tonsillectomy group (n = 119) | 6 episodes (SD 5.16) |
| Control group (n = 112) | 7.7 episodes (SD 5.88) |

B: Episodes of sore throat that invoked a GP consultation
The number of episodes of sore throat that evoked a GP consultation in the first year of the trial, <u>not including</u> the sore throat immediately following surgery, was as follows:

Adeno-/tonsillectomy group (n = 92)	1.9 episodes (SD 2.84)
Control group (n = 93)	2.4 episodes (SD 2.35)

The difference between the adenotonsillectomy group and the control treatment group in the number of episodes of sore throat that invoked a GP consultation is not statistically significant: -0.5 episodes (95% CI -1.2 to 0.3).

C: Sore throat days (including those immediately following surgery)

Data are reported on the number of days with sore throat including postoperative throat pain in the adeno-/tonsillectomy group:

Adeno-/tonsillectomy group (n = 110)	31.0 days (SD 29.25)
Control group (n = 107)	49.1 days (SD 54.62)

The difference between the adeno-/tonsillectomy group and the control treatment group is statistically significant: -18.1 days (95% CI -13.85 to -3.75).

Adverse events associated with surgery

Although surgical and anaesthetic morbidity was defined as an important secondary clinical outcome, no information on this endpoint is provided in the Lock 2010 publication.

Study 6: Alho 2007

The primary outcome measure in this trial was the proportion of patients with an acute episode of group A streptococcal pharyngitis during a 90-day follow-up period. Secondary outcomes included patient-recorded episodes (defined as at least two consecutive days) of sore throat and 'days with symptoms'. The secondary outcomes are reported as means and standard deviations "at the end of whole follow-up". The data reported relate to a "mean length of follow-up 164 days (SD 63) in control group and 170 days (SD 12) in tonsillectomy group".

The results of the Alho 2007 trial are based on tonsillectomy in 36 adults with recurrent "group A streptococcal pharyngitis" compared to 34 control subjects. In the control group, 5% (2/34) had elected to undergo surgery during follow-up, whereas all adults allocated to tonsillectomy received surgery. The results are presented as the mean number of episodes and sore throat days with accompanying standard deviations and by intention-to-treat.

A: Episodes of sore throat

For 'all episodes of pharyngitis' the reported data are:

Tonsillectomy group (n = 36)	0.6 episodes (SD 0.9) over follow-up period (mean 170 days)
Control group (n = 34)	2.1 episodes (SD 2.3) over follow-up period (mean 164 days)

The difference is reported as significant (P value = 0.001; Mann-Whitney U test). The data for the surgical group cannot include the immediate postoperative period as one episode of sore throat because the mean is less than one. The difference of 1.5 episodes (95% CI -2.33 to -0.67), occurring in the first five to six months after surgery, must therefore again be considered in the context of the one episode of postoperative pain.

B: Sore throat days (not including those immediately following surgery)

Data are reported on the number of days with sore throat "not including postoperative throat pain in tonsillectomy group":

Tonsillectomy group (n = 36)	3.2 days* (SD 5.3) over follow-up period (mean 170 days)
Control group (n = 34)	12.1 days (SD 14.1) over follow-up period (mean 164 days)

*_not including the 13 days due to surgery_
The difference is reported as significant (P value = 0.002; Mann-Whitney U test). In this study, the gain of 8.9 sore throat-free days (95% CI -13.95 to -3.85), occurring in the first five to six months after surgery, was bought at the cost of the days of postoperative pain. Patients in the tonsillectomy group had on average 13 days (SD 4) of throat pain during the immediate postoperative period.

Adverse events associated with surgery

There were no "serious adverse effects related to tonsillectomy" in any of the 70 adult participants. Two patients (6%) had secondary haemorrhages.

Study 7: Koskenkorva 2013

The primary outcome measure in this trial was the proportion of patients who had a severe episode of pharyngitis within five months. Secondary outcomes included patient-reported episodes (defined as at least two consecutive days) of sore throat, number of sore throat days and number of days absent from school and work. The secondary outcomes are reported as means and standard deviations "at the end of follow-up". The data reported relate to a "mean length of follow-up 5.7 months in the control group and 6.2 months in tonsillectomy group".

The results of the Koskenkorva 2013 trial are based on surgical intervention (tonsillectomy) in 46 adults with recurrent pharyngitis of any origin compared to 40 control subjects. In the control group, 8% (3/40) had elected to undergo surgery during follow-up, whereas all adults allocated to tonsillectomy received surgery. The results are presented as the mean number of episodes and sore throat days with accompanying standard deviations and by intention-to-treat.

A: Episodes of sore throat

For 'episodes of sore throat' the reported data are:

Tonsillectomy group (n = 46)	1.5 episodes (SD 2.4) over follow-up period (mean 6.2 months)
Control group (n = 40)	7.4 episodes (SD 5.8) over follow-up period (mean 5.7 months)

The difference is reported as significant (P value < 0.01; Mann-Whitney U test). The data for the surgical group did not include the immediate postoperative period. The difference of 5.9 episodes (95% CI -7.83 to -3.97), occurring in the first five to six months after surgery, must therefore again be considered in the context of the one episode of postoperative pain.

B: Sore throat days (NOT including those immediately following surgery)

Data are reported on the number of days with sore throat "not including postoperative throat pain in tonsillectomy group":

Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Review) 75

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

| Tonsillectomy group (n = 46) | 5.1* days (SD 8.7) over follow-up period (mean 6.2 months) |
| Control group (n = 40) | 40.2 days (SD 60.5) over follow-up period (mean 5.7 months) |

*_not including the 17 days due to surgery_
The difference is reported as significant (P value < 0.01; Mann-Whitney U test). In this study, the gain of 35.1 sore throat-free days (95% CI -54.02 to -16.18), occurring in the first five to six months after surgery, was bought at the cost of the days of postoperative pain. Patients in the tonsillectomy group had on average 17 days (SD 6) of throat pain during the immediate postoperative period.

Absence from school or work

Data are reported on the number of days absent from school or work "not including the postoperative period in tonsillectomy group":

| Tonsillectomy group (n = 46) | 3.3* days (SD 10.0) over follow-up period (mean 6.2 months) |
| Control group (n = 40) | 6.6 days (SD 11.8) over follow-up period (mean 5.7 months) |

*_not including the 14 days absence due to surgery_
The difference is reported as significant (P value = 0.02; Mann-Whitney U test). In this study, the gain of 3.3 days (95% CI -7.96 to 1.36), occurring in the first five to six months after surgery, was bought at the cost of the days absent from work or school immediately after surgery. Patients in the tonsillectomy group had on average 14 days (SD 6) absence from school or work during the immediate postoperative period.

Adverse events associated with surgery

Three patients (6%) were readmitted to the hospital; one patient (2%) because of severe throat pain and two (4%) because of mild secondary bleeding at one and four days after the operation.

F E E D B A C K

Paradise 2000

Summary

Editor's note: These comments relate to the original (1999) version of the review.
In a recent Cochrane review (Burton 1999) critiquing our randomized clinical trial of tonsillectomy in severely affected children (Paradise 1984), Burton, Towler, and Glasziou concluded that "significant baseline differences (in the history of antecedent throat infections and in parents' socioeconomic status) between the surgical and non-surgical groups and the inclusion of children who also underwent adenoidectomy prevent firm conclusions being drawn from the . . . trial." However, the Burton review fails to take into account a number of study features and findings that argue strongly against the importance of these factors as potential invalidators of our trial results.
First, consider the large differences in key outcomes favoring surgical over control subjects: in the first follow-up year a 14-fold reduction in throat infection episodes rated as moderate or severe (3 episodes in 38 surgical subjects vs 41 episodes in 35 control subjects), and in the second follow-up year, a 6-fold reduction (5 episodes in 31 surgical subjects vs 30 episodes in 29 control subjects). Other outcome differences were less dramatic but consistently in the same direction and also significant statistically.
Second, as we reported, tests for interaction albeit their limited power showed no significant differences in treatment outcomes that were related to any of the three factors cited by Burton et al (i.e. history of antecedent episodes, socioeconomic status, and presence or absence

of indications for adenoidectomy), nor were any of these factors related significantly to outcomes within the control group. Imbalances in factors that are not prognostic cannot fairly be considered sources of bias. Moreover, as we also reported, within each identifiable clinical and sociodemographic subgroup rates of throat infection were, without exception, lower for subjects treated surgically than for controls.

Third, consider the differences in antecedent history, which in any case may have been more apparent than real. Eligibility for our trial required a history of seven or more episodes of throat infection in the preceding year, five or more in each of the two preceding years, or three or more in each of the three preceding years. Not stated in our report were the facts that a number of children met more than one of these criteria and that such children were categorized as meeting the criterion involving the largest number of episodes. As chance would have it, more children in the surgical group than in the control group (20/43 vs 11/48) met the criterion of seven or more episodes in the preceding year. From this, Burton et al concluded that "the surgical group may therefore have included children with more severe disease," or "alternatively, these may have been children with less severe, but more short-lived disease." Setting aside for a moment that the analyses cited above argue against any prognostically important differences in disease severity, if the surgical group did indeed include children with more severe disease, the resulting bias would have favored control subjects rather than surgical subjects, in which case trial results would have understated, not overstated, the efficacy of surgery. Burton et al advance no rationale for their contrary, counterintuitive speculation that such children might actually have had less severe disease, but even if that had been the case the imbalance would hardly seem sufficient to account for the large differences in outcome.

Fourth, the difference in socioeconomic status referred to by Burton et al favored the control group rather than the surgical group. Again setting aside that the analyses cited above argued against the possibility that the difference was important prognostically, any resulting bias again might be expected to have favored control subjects, not surgical subjects.

In summary, to explain the large outcome differences we found favoring the surgical group on the basis of confounding would have required extreme imbalances between the surgical and control groups in variables that were strongly prognostic. In fact, however, not only were the variables of concern not apparently prognostic and their imbalances limited, but further, the expected effect of the imbalances would have been to favour the control group.

Finally, with regard to the adenoidectomy issue, Burton et al suggest that "Some part--potentially the greatest part--of the effect of 'surgery' could be due to removal of the adenoids." On the contrary--and again apart from the analyses described above--the addition of adenoidectomy in a minority of the surgically treated subjects could certainly not have accounted for the fact that moderate and severe throat-infection episodes were also virtually eliminated in the majority of such subjects who underwent tonsillectomy only.

We stand by our conclusion that in these severely affected children, tonsillectomy was unequivocally efficacious in reducing the occurrence of throat infection.

REFERENCES

Burton MJ, Towler B, Glasziou P. Tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Cochrane Review). In: The Cochrane Library, Issue 3, 1999. Oxford: Update Software.

Paradise JL, Bluestone CD, Bachman RZ, Colborn DK, Bernard BS, Taylor FH, Rogers KD, Schwarzbach RH, Stool SE, Friday GA, Smith IH, Saez CA. Efficacy of tonsillectomy for recurrent throat infection in severely affected children: Results of parallel randomized and nonrandomized clinical trials. N Engl J Med 1984;310:674-683.

Reply

Professor Paradise's comments have been addressed in this updated (2009) version of the review.

Contributors

Jack L Paradise
Children's Hospital of Pittsburgh
3705 Fifth Ave.
Pittsburgh
PA 15213
USA
Email: jpar+@pitt.edu

WHAT'S NEW

Date	Event	Description
3 November 2014	New citation required but conclusions have not changed	The general conclusions and recommendations regarding the effectiveness of adeno-/tonsillectomy versus (initial) non-surgical treatment remain unchanged We have reorganised the data-analysis, with the addition of new data since the last published update (Burton 2009). The data are now analysed as adeno-/tonsillectomy for adults and adeno-/tonsillectomy for children. We carried out additional sensitivity analyses to test whether any differences could be observed for adeno-tonsillectomy versus tonsillectomy alone, and to assess the impact of studies with high risk of attrition bias on the outcomes We have amended the analysis such that the data for children counted the episode of surgery as one episode of pain, and also counted the days of pain associated with the surgical procedure. However, it was not possible to do so for the data for adults, where the number of days of pain as a result of surgery were not included in the results for days with pain. Therefore, the data for adults still exclude the episode of pain and days of pain associated with the surgical procedure Two new review authors (Roderick Venekamp and Lee Yee Chong) joined the team to update this review
30 June 2014	New search has been performed	We updated the searches in June 2014. Three new trials have been included in the 2014 update (Koskenkorva 2013; Lock 2010; Stafford 1986). The first trial determined the effectiveness of tonsillectomy versus short courses of antibiotics for each episode of tonsillitis in 40 adults (Stafford 1986), but was at high risk of bias and did not provide any data suitable for analysis. The second trial determined the effectiveness of adeno-/tonsillectomy versus (initial) standard non-surgical treatment in children aged between 4 and 15 years with recurrent sore throats (four or more episodes of sore throat within each of two years or six or more episodes of sore throat within one year) and provided data on the mean number of 'any type of sore throat' and days with sore throat after one year of follow-up (Lock 2010). The third trial determined the short-term effectiveness of tonsillectomy versus watchful waiting in patients aged 13 years and older with recurrent pharyngitis of any origin (three or more episodes of pharyngitis in the previous 12 months) and provided data on the mean number of 'any type of sore throat', days with sore throat and days

(Continued)

		absent from school or work after five to six months follow-up (Koskenkorva 2013). We also identified one ongoing trial (NATTINA 2014).

HISTORY

Date	Event	Description
11 November 2008	New citation required and conclusions have changed	Review substantially updated following new searches in April 2008
11 April 2008	Amended	Converted to new review format.

CONTRIBUTIONS OF AUTHORS

Martin Burton: protocol development, searching for trials, quality assessment of trials, data extraction, review development.

Paul Glasziou: protocol and review development.

Roderick Venekamp: conducting searches, identifying studies, extracting data and preparing the manuscript for the updated review in 2014.

Lee Yee Chong: conducting searches, identifying studies, extracting data and preparing the manuscript for the updated review in 2014.

All authors have reviewed and provided comments on the updated version of the review.

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- None, Other.

External sources

- None, Other.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In this 2014 update, we have made a number of changes to the protocol to enhance clarity and to reflect important changes in methodology, views on the importance of outcomes and reporting format since the protocol was first published in 1999.

These include:

- inclusion of 'Summary of findings' tables;

- addition of further definitions of morbidity;

- including quality of life as a outcome to reflect the morbidity of the condition and the surgery;

- adding a line to clarify that the authors had an intention to analyse results from children and adults separately a priori;

- adding a sensitivity analysis by type of surgery (adenotonsillectomy or tonsillectomy).

NOTES

The original protocol for this review was titled 'Tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis'. The title was changed to include adenotonsillectomy when the first version of the full review was published. This reflected a change of scope which paralleled both day-to-day clinical practice and the interventions studied by the available studies. This is commented upon in the text of the review. The most important difference between the current version of the review and the previous one, Burton 2009, is the inclusion of three additional trials (Koskenkorva 2013; Lock 2010; Stafford 1986). The general conclusions and recommendations regarding the effects of adeno-/tonsillectomy versus (initial) non-surgical treatment in children and adults with chronic/recurrent acute tonsillitis on the number of any sore throat episodes, the number of moderate or severe sore throat episodes and the number of days with sore throat remained unchanged.

INDEX TERMS

Medical Subject Headings (MeSH)

*Adenoidectomy; *Tonsillectomy; Acute Disease; Chronic Disease; Pharyngitis [diagnosis]; Randomized Controlled Trials as Topic; Recurrence; Tonsillitis [surgery; *therapy]

MeSH check words

Adult; Child; Humans