Tonsillectomy: A Cost-Effective Option for Childhood Sore Throat? Further Analysis of a Randomized Controlled Trial

Janet A. Wilson, I. Nick Steen, Catherine A. Lock, Martin P. Eccles, Sean Carrie, Ray Clarke, Haytham Kubba, Chris H. Raine, Andrew Zarod and John Bond

Otolaryngology -- Head and Neck Surgery 2012 146: 122 originally published online 22 September 2011
DOI: 10.1177/0194599811422011

The online version of this article can be found at:
http://oto.sagepub.com/content/146/1/122

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
American Academy of Otolaryngology- Head and Neck Surgery

Additional services and information for Otolaryngology -- Head and Neck Surgery can be found at:

Email Alerts: http://oto.sagepub.com/cgi/alerts
Subscriptions: http://oto.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav

>> Version of Record - Dec 22, 2011
OnlineFirst Version of Record - Sep 22, 2011
What is This?
Tonsillectomy: A Cost-Effective Option for Childhood Sore Throat? Further Analysis of a Randomized Controlled Trial

Janet A. Wilson, FRCS¹,², I. Nick Steen, PhD¹, Catherine A. Lock, PhD¹, Martin P. Eccles, PhD¹, Sean Carrie, FRCS², Ray Clarke, FRCS³, Haytham Kubba, FRCS⁴, Chris H. Raine, FRCS⁵, Andrew Zarod, FRCS⁶, and John Bond, PhD¹

Abstract

Objective. To compare the estimated cost-effectiveness of childhood (adeno)tonsillectomy vs medical therapy for recurrent sore throats from the intention-to-treat (ITT) analysis of a randomized controlled trial (RCT) with that modeled on the recorded timing of surgical interventions as observed in all participants irrespective of their original group allocation.

Study Design. A pragmatic RCT (trial) with a parallel nonrandomized patient preference group (cohort) of (adeno)tonsillectomy vs medical therapy.

Setting. Five secondary care UK otolaryngology departments.

Subjects and Methods. Eligible children, aged 4 to 15 years, were enrolled to the trial (268) or cohort (461) groups. Outcomes included sore throat diaries, quality of life, and general practice consultations. The RCT protocol ITT analysis was compared with an as-treated analysis incorporating the cohort group, modeled to reflect the timing of tonsillectomy and the differential switch rates among the original groups.

Results. In the RCT ITT analysis, tonsillectomy saved 3.5 sore throats, whereas the as-treated model suggested an average reduction of more than 8 sore throats in 2 years for surgery within 10 weeks of consultation, falling to only 3.5 twelve months later due to the spontaneous improvement in the medical therapy group.

Conclusion. In eligible UK school-age children, tonsillectomy can save up to 8 sore throats at a reasonable cost, if performed promptly. Further prospective data collection, accounting for baseline and per-trial preferences and choice, is urgently needed.

Keywords

tonsillectomy, adenoidectomy, sore throat, pharyngitis, children
our intention-to-treat (ITT) analysis indicates that surgery saved 3.5 episodes of sore throat in 2 years. However, there are 3 major caveats with respect to this conclusion. First, as we had anticipated, is the high level of group switching, notably from the trial medical group to the surgical group. (One in 4 of those counted as nonsurgical in ITT terms actually had undergone surgery by the end of the follow-up period.) Second, our baseline analysis indicated that the children who consented to be randomized were not representative. Most severely affected children—in terms of long-lasting sore throats or time off school—elected to have surgery. The least severely affected appeared to opt to keep their tonsils. Only the “middle-ground” children were thus randomized. Finally, among those who did have surgery, the point in the 2-year review period at which they had surgery varied. Children who were randomized to tonsillectomy underwent tonsillectomy earlier in the follow-up observation of outcome period than did cohort children opting for surgery as parent/patient preference, whereas “switchers” were dispersed likewise. The aim of the present analysis is to remodel the findings to take account of these limitations in ITT (common to many surgical trials) and to reassess the clinical implications of the results in light of this “as actually treated” approach.

**Study Design**

The study design was a comprehensive cohort design comparing conventional surgical intervention with conventional medical treatment for children with recurrent sore throats. Eligible subjects who declined participation in a pragmatic, randomized controlled trial were offered their preferred treatment and invited to participate in the parallel, nonrandomized cohort study.

**Interventions**

Treatment consisted of tonsillectomy with or without adenoidectomy. Trial surgical interventions took place within 12 weeks of randomization, with preference for surgery according to local wait times at the participating hospitals. Patients in the nonsurgical groups were managed by their family doctor according to their usual practice.

**Participants, Inclusion Criteria, Randomization, and Sample Size**

Recruitment was from 2002 to 2006, with follow-up continuing until July 2008. Subjects were children aged 4 to 15 years on their last birthday with recurrent sore throats, referred to 1 of 5 secondary care otolaryngology departments in the northern United Kingdom by their family doctor. Initial inclusion criteria were a parental/older child account of ≥4 moderately severe or ≥6 mild episodes of sore throat per annum for 2 consecutive years; criteria were selected as the prevailing regional clinical guidelines for the performance of tonsillectomy. However, in the course of the trial, it became apparent that parents were unable to accurately recall the severity of individual episodes over 2 years in sufficient detail for this frequency metric to be workable. For instance, a parent citing 4 or 5 episodes could not say for certain that all 4 had lasted for more than 5 days. The criterion was amended to ≥5 episodes of sore throat for each of 2 years. Independent, Web-based randomization was stratified by hospital and by age group: 4 to 7 years, 8 to 11 years, and 12 to 15 years. The trial required 142 children in both study arms for a power of 80% to detect an effect size of 1 sore throat 2 years postrandomization, or 90% to detect a difference of 1.35 sore throats (α = 0.05).

**Outcome Measures and Data Collection**

Participants were followed up for 24 months from randomization (trial) or initial hospital consultation (cohort). The baseline data collection and findings are described elsewhere. In summary, 3 sets of data were collected:

1. Twenty-four 1-month structured sore throat health diaries, completed daily and returned monthly with a nonrespondent telephone reminder. An episode of sore throat was defined as a minimum of 3 consecutive days of sore throat. A gap of ≥4 non–sore throat days marked a new episode.
2. Questionnaires at 3, 12, and 24 months, including history of sore throat and associated problems with school, economic data, and the PedsQL Pediatric Quality of Life inventory, which has 2 generic core scales for physical and psychosocial health. Postal administration changed to telephone interview for the final 21 months of recruitment in an attempt to improve response rates.
3. Data on consultation rates and prescribed medication were sought from family doctor records at the end of follow-up for all trial participants and a sample of cohort participants.

For children <8 years, parents completed all outcomes. The between-group incremental cost-effectiveness ratio (ICER) was calculated by dividing the difference in costs between the 2 treatments by the difference in health effects.

**Statistical Methods**

The ITT analysis comprised analysis of (a) the number of episodes of sore throat in each of the 2 years of follow-up by a Poisson regression model with a log link function and an interval estimate of the difference in the mean number of sore throats over the 2-year follow-up, (b) primary care sore throat consultations using negative binomial regression, and (c) quality-of-life indices by independent sample t tests.

To address the ITT limitations, we set out to ascertain the difference in expected outcome (a) if all study children had had a tonsillectomy and (b) if all study children had received medical management. Because the characteristics of children who actually had a tonsillectomy differed from those who did not, it was not valid merely to compare children who had a tonsillectomy to those who did not. The following strategy was thus adopted:

1. What happened to children who did not have a tonsillectomy? How did the rate at which they reported...
sore throats vary with time? What factors influenced this rate? We developed a Poisson regression model to help us explain/describe the observed data.

2. We then repeated this process for those who did undergo tonsillectomy.

3. In each case, we explored how the frequency of reported sore throats varied with time and baseline characteristics of the child (such as the frequency of reported sore throats).

4. We then used these models to predict what would be expected to happen to the average child in the North of England and Scotland Study of Tonsillectomy and Adenotonsillectomy in Children (NESSTAC) (a) if he or she underwent tonsillectomy or (b) if his or her sore throats were managed medically.

The difference between these 2 scenarios presented us with our estimated effect of tonsillectomy, based on the assumption that the “average” NESSTAC child underwent that particular treatment.

The models are described in greater detail in the appendix (available at otojournal.org).

Confidence intervals for the predicted number of episodes of sore throat were derived using resampling procedures (bootstrapping)15 with 500 replications. The number of episodes of reported sore throat within a 4-weekly block was used as the unit of analysis. The 3 weeks following tonsillectomy were treated separately. A sore throat was allocated to the block in which the episode started.

Ethical Approval
This study was approved by the Northern & Yorkshire Multicentre Research Ethics Committee and associated local research ethics committees. The study received a Clinical Trial Authorisation from the Medicines Health Regulatory Authority. The study was approved by the NHS Research & Development department and Caldicott Guardian from each participating secondary care site. Primary Care Trust support was provided regarding collection of data from general practitioner records.

Results
Recruitment and Response Rates
We assessed 1546 children for eligibility, 729 of whom were enrolled: 268 were randomized in the trial and 461 entered the cohort. The mean diary return rate was similar in intervention arms (41%, trial; 29%, cohort participants; relative risk = 1.24; 95% confidence interval [CI], 1.16-1.81), with a decline in diary returns over the 24-month follow-up (Figure 1). The randomized groups were evenly balanced for age, sex, and frequency/severity of sore throats.12 Because of resource constraints, we could access 99% of trial subjects’ but only 48% of cohort subjects’ primary care records, which showed a mean of 9.8 consultations per participant per 2 years prestudy, with sore throat mentioned in 5.8. (This confirms that a significant minority of ear, nose, and throat clinic–reported sore throats appear not to have been the subject of prior family doctor consultation.)

Change of Group Status
Overall, 69 children (9.7%) switched groups, of whom 47 (17.5%) were from the trial arms (Figure 1). The highest switch rate was in the randomly allocated to medical management (26.3%) group.

ITT Analysis (Randomized [Trial] Participants Only), as Described in Full in the HTA Report
Over the 2-year period, the trial medical control group had a mean (SD) of 11.4 (8.7) episodes vs 7.4 (6.1) in those randomized to surgery (total days of sore throat are in Table 1). Poisson regression indicated that the incidence rate ratio (relative risk) in those randomized to the surgical group was 0.70 (95% CI, 0.61-0.80) in year 1 and 0.54 (0.42-0.70) in year 2. The overall reduction was 3.5 episodes over 2 years (95% CI, 1.8-5.2).11 Sore throat–related consultation rates were reduced by tonsillectomy, with an incidence rate ratio of 0.81 (95% CI, 0.59-1.10) in year 1 and 0.67 (0.46-0.97) in year 2. There were significant improvements in mean PedsQL14 quality of life for children in all 4 groups but no significant differences between groups at any time point (Table 2). In terms of costs, medical arm participants had more contact with primary care; surgery was associated with more hospital admissions. The estimated mean cost of medical management was £463.22, compared with £1402.15 for surgical management (2006 costs), with an incremental cost for surgery of £939 (ICER = £261 per each of the estimated 3.5 sore throats avoided; 95% CI, £161-£586).

“As-Treated” Analysis of the Study Total Population
The likelihood of reporting an episode of sore throat decreased over time. The 2 significant explanatory variables in the model were more frequent sore throats during the first 4 weeks and, likewise, having opted for surgery as a parent/patient preference. No additional explanatory variable significantly improved the fit to the observed data. However, it was also clear that the sooner the procedure is undertaken, the greater the benefit from tonsillectomy.

Table 3 indicates what happens to the “average” child over the 2 years following the initial consultation. In the first row, we assume that time to tonsillectomy for the average child is 4 weeks. Such children experience 0.7 episodes of sore throat during that 4-week wait. Children who are then managed nonsurgically for the remaining 100 weeks experience an average of another 12.3 sore throats vs 3.8 in those having tonsillectomy, a reduction of 8.5 (95% CI, 7.2-9.8) episodes over 2 years. In contrast, if the average time to tonsillectomy is taken as 1 year, only 3.5 (95% CI, 2.6-4.3) sore throats are avoided. Neither of these 2 estimates includes the sore throat due to the tonsillectomy operation. Table 3 indicates that the longer the tonsillectomy is delayed, the smaller the benefit. An approximation of the costs based on the treatments in trial participants—taking account of crossover...
between trial arms with an updated outcome of 8.21 fewer sore throats, again based on the incremental cost for surgery of £939 as above—gives clearly a much lower range of likely ICER per sore throat avoided.

**Discussion**

Tonsillectomy is among the most common childhood operations. This is the largest study to quantify the benefits of the
operation in children with moderate recurrent sore throat symptoms recruited across a range of centers serving ethnically diverse populations. Our preponderance of girls seeking secondary care opinion for sore throats, increasing with increasing age, has been noted in other countries. The results suggest that the effect of tonsillectomy in reducing the incidence of sore throat is larger in the as-treated analysis than in the ITT analysis. Having made the decision to operate, the sooner tonsillectomy is performed, the more cost-effective it appears to be. For children who have experienced at least 5 sore throats a year for 2 years, tonsillectomy after 10 weeks will prevent more than 8 sore throats compared with nonsurgical management over a 2-year period. Surgery deferred for a year, however, would prevent on average fewer than 4 sore throats. The report of 5+ annual sore throats exceeded the sore throat rates documented in general practitioner consultations—some might have presented to out-of-hours facilities, but most probably medical help may not have been sought. Parents nonetheless appear fairly sensitive to the diagnosis of tonsillitis. The per-protocol ITT analysis differed from the as-treated analysis due to the switching of a total of 47 of the randomized children (17.5%). The inclusion of the parent preference limbs highlighted the relative overall preference for surgery once referral to secondary care has occurred—of the 461 cohort participants, 86% had surgical management, suggesting a strong experiential societal belief in the effectiveness of tonsillectomy. This patient/parent consensus appears stronger than that of health care professionals.

The present study directly reflects current UK practice in 5 major centers and offers the best available evidence of the likely benefit of and natural history following tonsillectomy in children. Not all eligible patients were equally severely affected at baseline: only those whose sore throats were of intermediate severity agreed to be randomized. Overall switch rates in trials of childhood ear, nose, and throat surgery have been reported to be as high as 85%. The supplementary analysis highlights the degree to which the switch rate affected ITT analysis and allowed inclusion of the most and least severely affected, which allowed us to achieve our primary aim: to evaluate current practice pragmatically. We could map a trajectory of baseline severity from least (medical preference) through equipoise (randomized group) to most severe (surgical preference). One important flaw is the absence of accurate preenrollment data regarding the frequency of sore throat. However, our estimated baseline severity remained the most important predictor of clinical outcome. That is, parental recollection was at least

Table 1. Sore Throat Days and Related Health Care Costs over 2-Year Follow-up by Randomized Group

<table>
<thead>
<tr>
<th>Management</th>
<th>Medical Management (n = 115)</th>
<th>Surgical Management (n = 120)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of sore throat days per subject, year 1, mean (SD)</td>
<td>49.1 (54.6)</td>
<td>31.0 (29.3)</td>
</tr>
<tr>
<td>No. of sore throat days per subject, year 2, mean (SD)</td>
<td>20.2 (40.8)</td>
<td>8.0 (12.6)</td>
</tr>
<tr>
<td>Treatment for infection: total number of drugs prescribed (mean per patient [SD])</td>
<td>426 (5 [4.51])</td>
<td>389 (4 [3.77])</td>
</tr>
<tr>
<td>Analgesics: total number of drugs prescribed (mean number of drugs prescribed per patient [SD])</td>
<td>189 (2 [2.62])</td>
<td>211 (2 [3.13])</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total number of NHS contacts for sore throat–related consultations by all children over a 2-year follow-up</th>
<th>Medical Management</th>
<th>Surgical Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP contacts, including consultation, phone call, home visit</td>
<td>571</td>
<td>492</td>
</tr>
<tr>
<td>Nurse contact</td>
<td>106</td>
<td>84</td>
</tr>
<tr>
<td>Outpatient visits</td>
<td>79</td>
<td>64</td>
</tr>
<tr>
<td>Inpatient visits</td>
<td>15</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 2. PedsQL Subscale Scores for Physical Health and Psychosocial Health: Mean Scores at Baseline and Annual Rate of Change during the Period of the Study

<table>
<thead>
<tr>
<th>Management</th>
<th>Physical Health</th>
<th>Psychosocial Health</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline Mean</td>
<td>95% CI</td>
</tr>
<tr>
<td>Cohort medical</td>
<td>77.2</td>
<td>72.7</td>
</tr>
<tr>
<td>Cohort surgery</td>
<td>71.1</td>
<td>68.7</td>
</tr>
<tr>
<td>Trial medical</td>
<td>76.3</td>
<td>72.6</td>
</tr>
<tr>
<td>Trial surgery</td>
<td>78.8</td>
<td>75.4</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval. Quality of life scores are measured on a range of 1 to 100, with higher scores indicating better quality of life.
specific outcome measures such as the T14. We conclude that data might be aided by the use of recently developed disease-baseline severity assessment and completeness of outcome analysis. Benefits in much younger Dutch children errors of our results by as much as 50%, the results are still informative and potentially useful. Even if we assume that we have underestimated the standard trends and changes observed in the reporting of sore throat. However, there were large trends and changes observed in the reporting of sore throat. Even if we assume that we have underestimated the standard errors of our results by as much as 50%, the results are still informative and potentially useful.

Cost-effectiveness, of course, depends on the index population analyzed. Benefits in much younger Dutch children (aged 2-8 years) were not found to be cost-effective. In the United Kingdom, an estimated 75% of elective tonsillectomy procedures are for recurrent sore throats, and by 2008-2009, the number of tonsillectomies in children <15 years had fallen below 27,400 from almost 56,000 in 1994-1995. Conversely, the number of <15-year-olds admitted as emergencies to English hospitals with acute tonsillitis rose by 145% from below 27,400 from almost 56,000 in 1994-1995. Conversely, the number of <15-year-olds admitted as emergencies to English hospitals with acute tonsillitis rose by 145% from 6330 in 1999-2000 to 15,502 in 2008-2009. Policy makers for recurrent sore throats, and by 2008-2009, the number of tonsillectomies in children <15 years had fallen below 27,400 from almost 56,000 in 1994-1995. Conversely, the number of <15-year-olds admitted as emergencies to English hospitals with acute tonsillitis rose by 145% from 6330 in 1999-2000 to 15,502 in 2008-2009. Policy makers will be concerned to know whether these data are related. Although tonsillectomy for infrequent episodes of sore throat is uneconomic, for children with recurrent sore throats, the timing of surgery in relation to important events such as school exams may be critical. In our baseline analysis, concern about schooling was a significant predictor of preferring surgical treatment. The fact that there is a natural, untreated trend for symptoms to improve over time has been used as an argument for watchful waiting. But those with imminent, key ("once in a lifetime") educational milestones may regard the price of such a strategy as unacceptable. The concept of shared decision making offers a rational way forward—but in a publicly funded health care system such as in the United Kingdom, the lion’s share may rest with the policy maker rather than with the patient, the parent, or the otolaryngologist.

Although detailed economic evaluation was possible only for the ITT analysis, the dominance of the procedure itself in the costing means that our figure of £154 per sore throat saved represents a reasonably reliable estimate, provided that there is NHS capacity for surgery within 10 weeks of an initial consultation. For primary care clinicians, the present results provide an evidence base for assessing the appropriateness of tonsillectomy referral. Where children have at least a 2-year history of 5 annual sore throats, a delay of up to 12 months is counterproductive. The impact of a delay beyond 24 months remains unknown because of a lack of good data on the long-term natural history of childhood sore throat.

Acknowledgments

We sincerely thank all the children and their parents who took part in the study.

Author Contributions

Janet A. Wilson, lead clinical investigator, developed the protocol, was responsible for the acquisition of clinical data in secondary care, assisted in the interpretation of data, and drafted the manuscript; I. Nick Steen, trial statistician, developed the protocol, was responsible for the statistical analysis, prepared the results for publication, commented on drafts of the manuscript, and approved the final draft; Catherine A. Lock, trial manager and main researcher on the project, was responsible for the day-to-day management of the trial and the analysis of qualitative data, commented on drafts of the manuscript, and approved the final draft; Martin P. Eccles, developed the protocol, was responsible for the acquisition of clinical data from general practice and assisted in the analysis and interpretation of data.

---

Table 3. Estimated Number of Episodes of Sore Throat Saved in 2 Years from Initial Consultation

<table>
<thead>
<tr>
<th>Time</th>
<th>Medical Management</th>
<th>Tonsillectomy</th>
<th>Sore Throats Saveda</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n0</td>
<td>n 95% CI</td>
<td>n 95% CI</td>
</tr>
<tr>
<td>4</td>
<td>0.7</td>
<td>12.3</td>
<td>11.1</td>
</tr>
<tr>
<td>6</td>
<td>1.1</td>
<td>11.9</td>
<td>10.8</td>
</tr>
<tr>
<td>8</td>
<td>1.5</td>
<td>11.6</td>
<td>10.4</td>
</tr>
<tr>
<td>10</td>
<td>1.8</td>
<td>11.2</td>
<td>10.0</td>
</tr>
<tr>
<td>12</td>
<td>2.2</td>
<td>10.9</td>
<td>9.7</td>
</tr>
<tr>
<td>16</td>
<td>2.8</td>
<td>10.2</td>
<td>9.1</td>
</tr>
<tr>
<td>26</td>
<td>4.4</td>
<td>8.6</td>
<td>7.5</td>
</tr>
<tr>
<td>39</td>
<td>6.3</td>
<td>6.8</td>
<td>5.7</td>
</tr>
<tr>
<td>52</td>
<td>7.9</td>
<td>5.1</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval. Time = time in weeks from initial consultation to tonsillectomy. n0 = estimated number of episodes of sore throat between the consultation and the tonsillectomy. n = estimated number of episodes of sore throat during the remaining period of follow-up. Note the number of episodes of sore throats saved does not take into account any episodes of sore throat associated with the actual tonsillectomy procedure.
commented on drafts of the manuscript, and approved the final draft; **Sean Carrie**, recruited participants and assisted in the acquisition and interpretation of data, commented on drafts of the manuscript, and approved the final draft; **Ray Clarke**, recruited participants, assisted in the acquisition and interpretation of data, and approved the final draft; **Haytham Kubba**, recruited participants, assisted in the acquisition and interpretation of data, and approved the final draft; **Chris H. Raine**, recruited participants, assisted in the acquisition and interpretation of data, and approved the manuscript; **Andrew Zarod**, recruited participants, assisted in the acquisition and interpretation of data, and approved the final draft; **John Bond**, the lead investigator, was responsible for the overall management of the study, developed the protocol, and assisted in the preparation of the manuscript.

**Disclosures**

**Competing interests:** None.

**Sponsorships:** Newcastle upon Tyne Hospitals NHS Foundation Trust. Sponsor charged with research governance, good clinical practice, Caldicott compliance only.

**Funding source:** The UK Department of Health through the National Institute for Health Research, Health Technology Assessment Programme (grant number 99/20/03).

**Supplemental Material**

Additional supporting information may be found at http://oto.sagepub.com/content/by/supplemental-data

**References**