

# Conservative management versus tonsillectomy in adults with recurrent acute tonsillitis in the UK (NATTINA): a multicentre, open-label, randomised controlled trial

Janet A Wilson\*, James O'Hara\*, Tony Fouweather, Tara Homer, Deborah D Stocken, Luke Vale, Catherine Haighton, Nikki Rousseau, Rebecca Wilson, Lorraine McSweeney, Scott Wilkes, Jill Morrison, Kenneth MacKenzie, Kim Ah-See, Sean Carrie, Claire Hopkins, Nicola Howe, Musheer Hussain, Hisham Mehanna, Christopher Raine, Frank Sullivan, Alexander von Wilamowitz-Moellendorff, M Dawn Teare



## Summary

**Background** Tonsillectomy is regularly performed in adults with acute tonsillitis, but with scarce evidence. A reduction in tonsillectomies has coincided with an increase in acute adult hospitalisation for tonsillitis complications. We aimed to assess the clinical effectiveness and cost-effectiveness of conservative management versus tonsillectomy in patients with recurrent acute tonsillitis.

**Methods** This pragmatic multicentre, open-label, randomised controlled trial was conducted in 27 hospitals in the UK. Participants were adults aged 16 years or older who were newly referred to secondary care otolaryngology clinics with recurrent acute tonsillitis. Patients were randomly assigned (1:1) to receive tonsillectomy or conservative management using random permuted blocks of variable length. Stratification by recruiting centre and baseline symptom severity was assessed using the Tonsil Outcome Inventory-14 score (categories defined as mild 0–35, moderate 36–48, or severe 49–70). Participants in the tonsillectomy group received elective surgery to dissect the palatine tonsils within 8 weeks after random assignment and those in the conservative management group received standard non-surgical care during 24 months. The primary outcome was the number of sore throat days collected during 24 months after random assignment, reported once per week with a text message. The primary analysis was done in the intention-to-treat (ITT) population. This study is registered with the ISRCTN registry, 55284102.

**Findings** Between May 11, 2015, and April 30, 2018, 4165 participants with recurrent acute tonsillitis were assessed for eligibility and 3712 were excluded. 453 eligible participants were randomly assigned (233 in the immediate tonsillectomy group vs 220 in the conservative management group). 429 (95%) patients were included in the primary ITT analysis (224 vs 205). The median age of participants was 23 years (IQR 19–30), with 355 (78%) females and 97 (21%) males. Most participants were White (407 [90%]). Participants in the immediate tonsillectomy group had fewer days of sore throat during 24 months than those in the conservative management group (median 23 days [IQR 11–46] vs 30 days [14–65]). After adjustment for site and baseline severity, the incident rate ratio of total sore throat days in the immediate tonsillectomy group (n=224) compared with the conservative management group (n=205) was 0.53 (95% CI 0.43 to 0.65; <0.0001). 191 adverse events in 90 (39%) of 231 participants were deemed related to tonsillectomy. The most common adverse event was bleeding (54 events in 44 [19%] participants). No deaths occurred during the study.

**Interpretation** Compared with conservative management, immediate tonsillectomy is clinically effective and cost-effective in adults with recurrent acute tonsillitis.

**Funding** National Institute for Health Research.

**Copyright** © 2023 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

## Introduction

Tonsillectomy for patients with recurrent acute tonsillitis is one of the most common adult operations performed by the National Health Service (NHS) in England, with approximately 16 000 procedures per year.<sup>1</sup> In the USA, 102 000 tonsillectomy procedures are done annually.<sup>2</sup> The health-care cost of sore throat episodes in the UK is estimated to be as high as £2.35 billion (US\$3.20 billion).<sup>3</sup> In the past 20 years, tonsillectomy rates have reduced by up to 50% in many European countries, whereas hospital admissions for tonsillitis have increased by 136%.<sup>4</sup>

Tonsillectomy is a painful procedure which requires an average of 14 days off work.<sup>5,6</sup> The level 1 evidence for tonsillectomy in adults is scarce,<sup>5</sup> which probably contributes to variation in tonsillectomy rates between countries.<sup>7</sup> The 2014 Cochrane Review identified two studies with 156 participants and concluded that the evidence for tonsillectomy in adults was of low quality.<sup>5</sup> A meta-analysis showed that the number of sore throat days was 10.6 days (95% CI 5.8–15.8) fewer in patients receiving tonsillectomy than those treated conservatively during approximately 6 months of follow-up. However,

Published Online  
May 17, 2023  
[https://doi.org/10.1016/S0140-6736\(23\)00519-6](https://doi.org/10.1016/S0140-6736(23)00519-6)

See Online/Comment  
[https://doi.org/10.1016/S0140-6736\(23\)00673-6](https://doi.org/10.1016/S0140-6736(23)00673-6)

\*Joint first authors

Population Health Sciences Institute (Prof J A Wilson MD, J O'Hara MD, T Fouweather BSc, T Homer MSc, Prof L Vale PhD, L McSweeney PhD, Prof M D Teare PhD) and Newcastle Clinical Trials Unit (R Wilson MSc, N Howe PhD, A von Wilamowitz-Moellendorff PhD), Newcastle University, Newcastle upon Tyne, UK; Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK

(Prof D D Stocken PhD, N Rousseau PhD); Department of Social Work, Education, and Community Wellbeing, Northumbria University, Newcastle upon Tyne, UK (Prof C Haighton PhD); Ear, Nose, and Throat Department, Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK (J O'Hara, S Carrie MB ChB); School of Medicine Faculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, UK

(Prof S Wilkes PhD); Senate Office, University of Glasgow, Glasgow, UK (Prof J Morrison PhD); Department of Ear, Nose, and Throat Surgery, NHS Greater Glasgow and Clyde, Glasgow, UK (Prof K MacKenzie FRCS); Department of Otolaryngology Head and Neck Surgery, NHS Grampian, Aberdeen, UK (Prof K Ah-See MD); Ear, Nose and Throat Department and Head and Neck Department, Guy's and St Thomas' NHS Foundation Trust, London, UK (Prof C Hopkins DM Oxon); School of Medicine, Dundee

University, Dundee, UK  
(Prof M Hussain MBBS);  
Institute of Head and Neck  
Studies and Education,  
University of Birmingham,  
Birmingham, UK  
(Prof H Mehanna PhD); Ear,  
Nose, and Throat Department,  
Bradford Teaching Hospitals  
NHS Foundation Trust,  
Bradford, UK  
(Prof C Raine ChM); Population  
and Behavioural Science  
Division, School of Medicine,  
University of St Andrews,  
St Andrews, UK  
(Prof F Sullivan FRSE)

Correspondence to:  
Mr James O'Hara, Population  
Health Sciences Institute,  
Newcastle University, Newcastle  
upon Tyne NE2 4AX, UK  
james.o'hara@newcastle.ac.uk

### Research in context

#### Evidence before this study

Iterations of the Cochrane Review on the effectiveness of tonsillectomy in adults with recurrent tonsillitis identified low-quality evidence. Two small, randomised trials with short follow-up suggested that the number of sore throat days might be fewer in the first 6 months after a tonsillectomy than after conservative management. Given scarce evidence to support this common intervention, the UK National Institute for Health and Care Research commissioned a pragmatic clinical trial. In line with Cochrane Review recommendations, the brief specified that the primary outcome should be sore throat days during 24 months. We searched MEDLINE for clinical trials published between Jan 1, 1996, and Jan 1, 2022, using the search term "tonsillectomy", limited to randomised trials and adults and found no further relevant clinical trials.

#### Added value of this study

To our knowledge, NATTINA is the largest randomised trial to assess the clinical effectiveness and cost-effectiveness of a surgical intervention in adults with a sore throat. We showed that compared with conventional medical management, tonsillectomy is clinically effective and cost-effective. The incident rate ratio of sore throat days was 0.53 with tonsillectomy during 24 months on the most conservative estimate and 0.42 when accounting for actual treatment administered.

#### Implications of all the available evidence

NATTINA quantifies the comparative value of tonsillectomy for health service providers, optimises the basis for shared decision making in adults with recurrent tonsillitis, and thus, has the potential to reduce regional variation in referrals for tonsillectomy. But predominantly, NATTINA encourages timely access to care.

neither study accounted for postoperative sore throat days. NATTINA was commissioned by the NHS National Institute for Health and Care Research (NIHR) with an aim of addressing the evidence gap by assessing the clinical effectiveness and cost-effectiveness of conservative management versus tonsillectomy in patients with recurrent acute tonsillitis.

## Methods

### Study design and participants

This pragmatic multicentre, open-label, randomised controlled trial was conducted in 27 hospitals in the UK. Participants were adults aged 16 years or older who were newly referred to secondary care otolaryngology clinics with recurrent acute tonsillitis. Recruited participants had to meet the UK guidelines for tonsillectomy, which stipulate that sore throat episodes are due to acute tonsillitis; the episodes of sore throat prevent healthy functioning; and that there are seven or more clinically significant sore throat episodes in the preceding year, or five or more episodes in each of the preceding 2 years, or three or more episodes in each of the preceding 3 years.<sup>8</sup> We did not aim to address the evidence rationale for these largely consensus-based guidelines. The full list of inclusion and exclusion criteria is shown in the appendix (p 1). All participants provided written informed consent. The trial protocol<sup>9</sup> was approved by the North East Research Ethics Committee (November, 2014; 14/NE/1144). Oversight was done by independent data monitoring and trial steering committees.

### Randomisation and masking

Patients were randomly assigned (1:1) to receive tonsillectomy or conservative management using random permuted blocks of variable length. Randomisation was performed centrally by the Newcastle Clinical Trials Unit Randomisation Service, which is an in-house, bespoke,

internet-based system. A statistician (Newcastle University, Newcastle, UK), otherwise not involved in the trial, produced the final allocation schedule. Stratification by recruiting centre and baseline symptom severity was assessed using the Tonsil Outcome Inventory-14 (TOI-14) score (categories defined as mild 0–35, moderate 36–48, or severe 49–70).<sup>10</sup> No masking was done in this study.

### Procedures

To reflect the pragmatic nature of this trial because tonsillectomy is a commonly performed operation by surgeons, including supervised trainee surgeons, the recruiting centre was used as a marker of the surgical team. Data on individual surgeons were not collected. Participants in the tonsillectomy group received elective surgery to dissect the palatine tonsils within 8 weeks after random assignment. Surgical methods used were the same as in standard care at recruitment sites and no stipulations were made by the trial team regarding the choice of surgical approach. Participants in the conservative management group received standard non-surgical care during 24 months, comprising of self-administered analgesia plus ad hoc primary care prescription of antibiotics or attendance at emergency departments for severe tonsillitis.

For the primary outcome, once per week data were collected mostly with a text message (email and telephone calls were also available; return of data defined as returns). To reflect routine practice and national guidelines, and maximise symptom data capture, throat swabs were not required to report a sore throat. Surgical outcome follow-up after a tonsillectomy occurred by telephone 1–2 weeks after the surgery.

For secondary outcomes, questionnaires (TOI-14 and 12-Item Short Form Survey [SF-12]) were posted at 6 and 18 months after random assignment and participants followed up in person at 12 and 24 months. TOI-14 is a

See Online for appendix

patient-reported, disease-specific, quality of life (QoL) questionnaire reflecting symptoms in the previous 6 months, validated in patients with chronic tonsillitis who had a tonsillectomy. Each item is a 0–5 Likert score with an adjusted total score of 100 (higher scores indicate worse QoL). The SF-12 is a shortened version of the SF-36, comprised of SF-12 Mental Component Summary (MCS) scores and SF-12 Physical Component Summary (PCS) scores.<sup>11</sup> The maximum score for both components is 100 (higher scores indicate better QoL).

### Outcomes

The primary outcome was the number of sore throat days collected during 24 months after random assignment. Secondary outcomes were TOI-14 score at baseline and at 6, 12, 18, and 24 months; general QoL measured using the SF-12 at baseline and at 6, 12, 18, and 24 months; the number of adverse events; and the views and experiences of patients and clinicians regarding tonsillectomy and conservative management (reported elsewhere).<sup>12</sup>

Economic evaluation outcomes (part of secondary outcomes) were the costs incurred by health-care providers to manage reoccurring sore throat episodes, collected with case report forms and directly from participants using bespoke Health Utilisation Questionnaires (HUQs) at baseline and 6, 12, 18, and 24 months (including a time and travel questionnaire at 18 months); direct and indirect costs incurred by participants collected with self-reported questionnaires; quality-adjusted life-years (QALYs) based on responses to the SF-12 mapped onto the Short Form 6-Dimension (SF-6D) instrument;<sup>13</sup> and incremental cost per QALY gained.

### Statistical analysis

The sample size of 510 participants was based on 90% power to detect a standardised effect size of 0.33 (mean sore throat intergroup difference of 3.6 days [pooled estimated SD 10.8]) in a quantitative outcome, assuming a type 1 error rate of 5% (appendix p 1), allowing for 20% loss to follow-up and 5% crossovers (ie, those who switched treatment groups). This effect size was based on previous research.<sup>14</sup> Due to slower than expected recruitment, a substantial amendment was approved by the independent trial steering committee and funder, to reduce the power to 85% and recruitment sample size to 444 (retaining the same detectable difference).

The primary analysis was done in the intention-to-treat (ITT) population (defined as randomly assigned participants who returned primary outcome data) using mixed-effect multivariable negative binomial regression with outcome variable sore throat days, adjusted for the recruiting centre (as a random effect) and baseline severity (as a fixed effect). A two-sided *p* value of less than 0.05 was considered as significant.

To account for missing data during 104 weeks of collection, the rate of return of sore throat data was

included as an exposure variable in the negative binomial regression model. We assumed that these once per week data were missing at random. The effect of randomisation is summarised with the incident rate ratio (defined as the ratio of predicted numbers of sore throat days in the two treatment groups).

Pre-planned primary sensitivity analyses included a per protocol analysis restricted to participants who had a tonsillectomy within 8 weeks after random assignment compared with those who remained within the conservative group and did not cross over to surgery; and a per treatment analysis of participants who received a tonsillectomy at any time during follow-up compared with those who received conservative management, regardless of the initial randomisation group.

Two unplanned sensitivity analyses were performed. The first analysis assumed that a missing sore throat return meant that there were no sore throat episodes to report that week, which was done by omitting the exposure variable in the negative binomial regression model for the ITT analysis. The second repeated the primary ITT analysis for participants who returned at least 80% of the sore throat data.

TOI-14 scores at each timepoint were compared using descriptive statistics and maximum likelihood computations with repeated measures, adjusted for treatment groups and stratification factors. A two-sample *t* test quantified the between-group difference at 12 and 24 months. SF-12 scores were analysed using repeated measures in a similar way to TOI-14 scores. Data were processed using the Optum PRO CoRE software (SF-12 version 2) to derive the MCS and PCS scores. In the economic evaluation, primary and secondary health-care costs were based on cost of surgery and self-reported health-care use, collected with the HUQ and combined with unit costs from routine sources.<sup>15</sup> Total health-care cost per participant was estimated and summarised as the average total cost per group.

A time and travel questionnaire administered at 18 months collected data on direct (eg, parking) and indirect (eg, time away from usual activities) costs incurred by participants to attend the health-care appointments and was used to derive a standard set of unit costs for time and travel costs. The questionnaire was used for the first 2 years of recruitment and stopped thereafter to reduce participant response burden. Direct (eg, over-the-counter medication) and indirect (eg, time away from usual activities) costs associated with sore throat episodes were collected with the bespoke HUQ questionnaire administered when a participant reported a sore throat in the previous week. These data were collated when participants reported a health-care contact using the HUQ questionnaire or a sore throat day with return of sore throat data.

Utility values from the SF-6D were used to estimate QALYs using the area under the curve.<sup>16</sup> Costs and QALYs incurred in the second year of follow-up were

For the Optum PRO CoRE software see <https://www.qualitymetric.com/>

discounted at 3·5%.<sup>17</sup> Missing cost and utility data were assumed to be missing at random. Chained multiple imputation by predictive mean matching was used for missing data.<sup>18</sup> Missing data were imputed simultaneously using regression models controlled for baseline characteristics (TOI-14, age, sex, and ethnicity). Differences in average costs were estimated using seemingly unrelated regression.<sup>19</sup> Uncertainty in results was estimated using stochastic sensitivity analysis (bootstrapping).<sup>20</sup> The bootstrapping method estimates the difference in costs and QALYs between two participants (one from each group) with resampling, and this method was done for 1000 samples. A cost-effectiveness acceptability curve and plane were used to illustrate uncertainty in the estimation.<sup>21,22</sup> Cost-effectiveness acceptability curves were replicated with total costs converted into US\$ and € using

the Cochrane Economic Methods Group Cost Converter.<sup>23</sup> The economic analysis was done following best practice guidelines.<sup>24</sup>

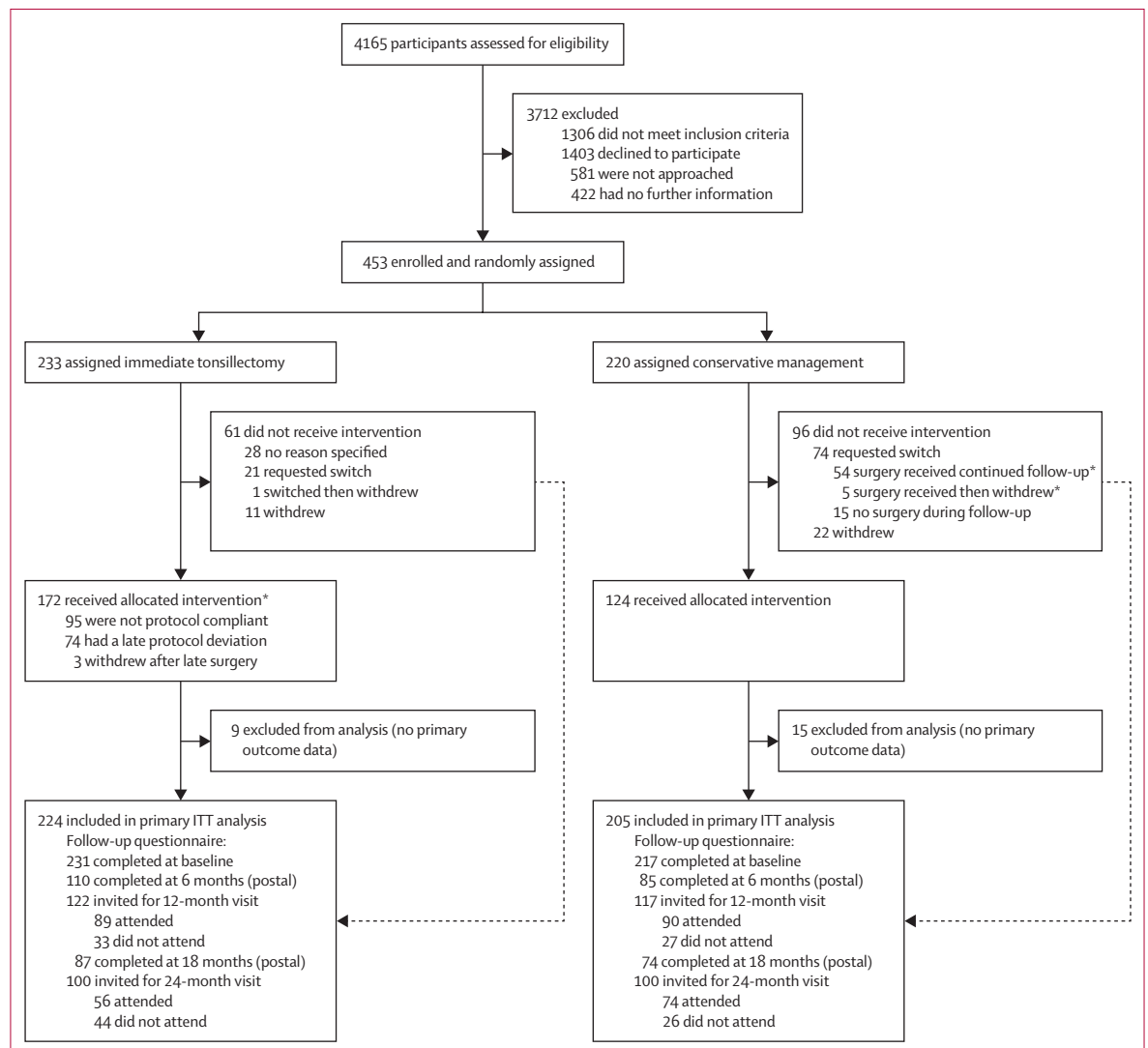
All statistical analyses were performed using STATA (version 16). This study is registered with the ISRCTN registry, 55284102.

### Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

### Results

Between May 11, 2015, and April 30, 2018, 4165 participants with recurrent acute tonsillitis were assessed for eligibility and 3712 were excluded (figure 1). 453 eligible



**Figure 1: Trial profile**

Numbers at the allocation stage differ to the numbers in the final ITT, per protocol, and per treatment analyses due to an absence of once per week sore throat returns (weeks 1-105; primary outcome). ITT=intention-to-treat. \*In total, 231 underwent tonsillectomy.

participants were randomly assigned (233 in the immediate tonsillectomy group and 220 in the conservative management group). Of those, 157 (35%) did not receive the allocated intervention. 429 (95%) patients were included in the primary ITT analysis (224 in the tonsillectomy group and 205 in the conservative management group). 24 (5%) were excluded from the analysis because they did not have primary outcome data.

Baseline demographics were well balanced across treatment groups (table 1). The median age of participants was 23 years (IQR 19–30), with 355 (78%) females and 97 (21%) males. Most participants were White (407 [90%]). Univariate and subsequent multivariate analysis showed that sex and educational levels were not independently associated with treatment outcomes. TOI-14 was collected for 493 (13%) of 3712 patients who declined entering the trial. This group had higher baseline TOI-14 scores than those entering the trial (appendix p 2).

Of 231 participants who underwent tonsillectomy, 122 (53%) were cold dissection, 91 (39%) were bipolar diathermy, one (<1%) was coblation, one (<1%) was laser, eight (3%) were mixed (cold and bipolar dissection), and one (<1%) was done outside of the trial site. Details of tonsillectomy type were not received for seven (3%) patients. Grade of surgeon experience was recorded for 219 tonsillectomies (116 [53%] were performed by consultants, 75 [34%] by higher surgical trainees of otolaryngology, 15 [7%] by non-trainee otolaryngology surgeons, and 13 [6%] by junior otolaryngology trainees or other grade of surgeon; data not shown).

Complete sore throat data for each of the 104 weeks was received for 115 (25%) of 453 participants. 15 (3%) did not provide sore throat data. There was a numerically higher return rate of sore throat data in the tonsillectomy group than in the conservative management group (144 [62%] vs 115 [53%] patients completed sore throat data returns for 83 [80%] of 104 weeks). Cumulative return rates are shown in the appendix (pp 8–9).

Of 429 patients, data collected in 42 (10%) were retained in the analysis until the point of withdrawal. Participants in the immediate tonsillectomy group had fewer days of sore throat during 24 months than those in the conservative management group (median 23 days [IQR 11–46] vs 30 days [14–65]; figure 2).

After adjustment for site and baseline severity, the incident rate ratio of total sore throat days in the immediate tonsillectomy group (n=224) compared with the conservative management group (n=205) was 0.53 (95% CI 0.43–0.65; p<0.0001; table 2). The grade of surgeon experience was not statistically significant with negative binomial multivariable regression (appendix pp 9–11). Figure 2 shows ITT incident rate ratio on a log scale with the results of planned sensitivity analyses in the per protocol and per treatment populations and two unplanned sensitivity analyses. The planned and unplanned sensitivity analyses supported the primary

	Immediate tonsillectomy (n=233)	Conservative management (n=220)
<b>Sex</b>		
Female	175 (75%)	180 (82%)
Male	57 (24%)	40 (18%)
Missing*	1 (<1%)	0
Median age, years	23 (19–30)	23 (19–30)
<b>Ethnicity</b>		
White	211 (91%)	196 (89%)
Asian (Indian, Pakistani, or Bangladeshi)	9 (4%)	9 (4%)
Asian other	0	4 (2%)
Black or African Caribbean	6 (3%)	6 (3%)
Other	6 (3%)	5 (2%)
Missing*	1 (<1%)	0
<b>Education level</b>		
Postgraduate	20 (9%)	23 (10%)
Degree, professional, or vocational	48 (21%)	70 (32%)
Higher, A-level, national grade, or vocational	100 (43%)	74 (34%)
O-level, O grade, GCSE, standard grade, or vocational	57 (24%)	47 (21%)
No qualification	5 (2%)	1 (<1%)
Missing*	3 (1%)	5 (2%)
<b>Employment status</b>		
Self-employed	12 (5%)	11 (5%)
Employed (full or part time)	149 (64%)	132 (60%)
Unemployed (actively seeking work)	6 (3%)	8 (4%)
Full time student or at school	48 (21%)	52 (24%)
Other†	15 (6%)	13 (6%)
Missing*	3 (1%)	4 (2%)
<b>Tonsil Outcome Inventory-14 score</b>		
Median	44 (36–51)	44 (37–51)
Missing	2 (1%)	3 (1%)
<b>Symptom severity</b>		
Mild (0–35)	51 (22%)	44 (20%)
Moderate (36–48)	94 (40%)	96 (44%)
Severe (49–70)	88 (38%)	80 (36%)
<b>SF-12 Mental Component Summary score</b>		
Median	47 (39–54)	47 (37–55)
Missing	2 (1%)	4 (2%)
<b>SF-12 Physical Component Summary score</b>		
Median	49 (45–55)	51 (45–56)
Missing	2 (1%)	4 (2%)

Data are n (%) or median (IQR). SF-12=12-Item Short Form Survey. \*Data for one participant could not be recovered from the recruiting centre. †Retired, on maternity leave, looking after family at home, disabled, with long-term sickness, or receiving a government training scheme.

**Table 1: Baseline characteristics**

ITT analysis with fewer sore throat days reported in participants who were randomly assigned or received tonsillectomy. The per protocol analysis showed a greater reduction in sore throat days in the tonsillectomy group

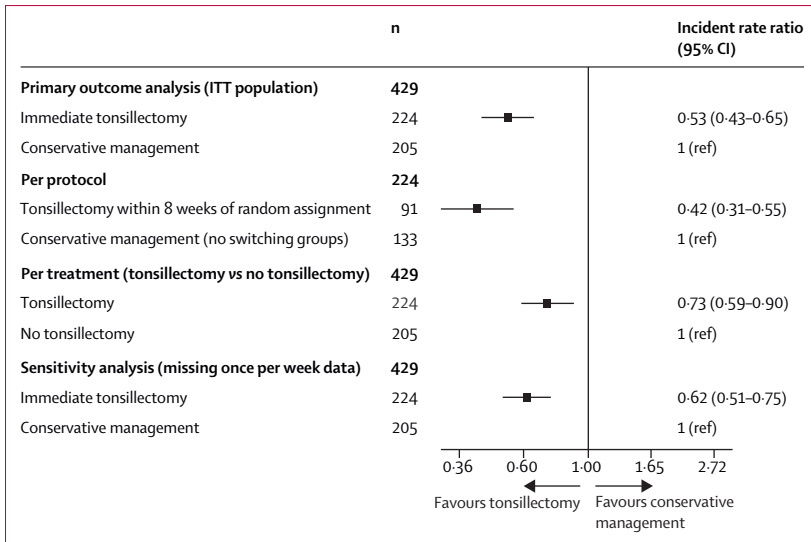


Figure 2: Forest plots of primary outcome analyses  
ITT=intention-to-treat.

	Incidence rate ratio (95% CI)	p value
<b>Treatment groups</b>		
Conservative management	1 (ref)	..
Immediate tonsillectomy	0.53 (0.43 to 0.65)	<0.0001
<b>Baseline severity</b>		
Mild	1 (ref)	..
Moderate	1.21 (0.92 to 1.59)	0.17
Severe	1.03 (0.78 to 1.37)	0.83
Constant	122.48 (87.52 to 171.40)	<0.0001
Natural log (exposure)	1	..
Natural log of $\alpha$	0.03 (-0.10 to 0.15)	..
$\alpha$ (dispersion parameter)	1.03 (0.90 to 1.17)	..

Log likelihood (-2074.77) and likelihood ratio test versus negative binomial model (chibar2=17.12). Prob>chibar2<0.0001. The variance parameter for random variable site is 0.19 (95% CI 0.06-0.55), chibar2 equals  $\chi^2$  with 1 degree of freedom; the test is a restricted likelihood ratio for variance. The Kenward-Roger approximation was used to estimate the denominator degrees of freedom.

**Table 2: Multilevel mixed effect negative binomial regression (intention-to-treat population; n=429)**

compared with conservative management (incident rate ratio 0.42 [95% CI 0.31-0.55]). The per treatment analysis shows a smaller reduction in sore throat days in the tonsillectomy group compared with no tonsillectomy (incident rate ratio 0.73 [0.59-0.90]). The mean sore throat data in the two study groups and for participants crossing over from randomly assigned groups are shown in the appendix (p 9).

Baseline TOI-14 data were completed by 448 (99%) of 453 participants and SF-12 data were completed by 447 (99%) participants. Completion rates were 239 (53%) for both at the 12-month visit and 199 (44%) for TOI-14 and 200 (44%) for SF-12 at the 24-month visit. TOI-14 scores were similar for both groups at baseline

(table 1). Figure 3 shows that TOI-14 scores improved (reduced) during 24 months in both treatment groups, with more pronounced and earlier improvement in the immediate tonsillectomy than in the conservative management group (mean 4.3 [95% CI 3.4-5.3] in 122 [55%] participants vs 21.7 [18.5-24.9] in 117 [50%] at 12 months and 4.7 [2.9-6.4] in 99 [42%] vs 15.4 [12.0-18.8] in 100 [45%] at 24 months; p<0.0001). SF-12 scores were also higher during 24 months of follow-up (better QoL) in the immediate tonsillectomy group (appendix pp 13-15).

Responses to the HUQ were near complete at baseline, with 231 (99%) participants responding in the immediate tonsillectomy group versus 215 (98%) in the conservative management group, which reduced to 122 (52%) versus 117 (53%) at 12 months and to 99 (42%) versus 100 (45%) at 24 months. Details of health-care resource use by study groups are shown in the appendix (pp 15-20). Based on imputed results, tonsillectomy was non-significantly more costly (mean difference £488 [95% CI 349-626]) and more effective (mean difference 0.12 QALYs [0.09-0.14]) than conservative management. The incremental cost per QALY gained was £4136 (appendix p 23). Results from the stochastic analysis are presented in the appendix (pp 24-26). Over the range of values considered for an additional QALY, tonsillectomy has a high probability of being considered cost-effective. Assuming a £5000 (\$5000 or €5000) threshold value for an additional QALY, tonsillectomy had an 85% (17% with \$5000 and 63% with €5000) probability of being considered cost-effective. This finding increased to 100% when an additional QALY was valued at £10000 (\$10000 or €10000). When participant costs were considered, tonsillectomy was less costly than conservative management (mean difference £889 [95% CI 40-1738]) and more effective (mean difference 0.12 QALYs [0.09-0.14]).

There were 191 adverse events in 90 (39%) of 231 participants undergoing tonsillectomy that were deemed related to tonsillectomy (table 3). Serious adverse events included bleeding (37 events in 34 [15%] participants) and infection (six events in six [3%]). The most common adverse event was bleeding (54 events in 44 [19%] participants). No deaths occurred during the study. There were a further 15 severe adverse events reported that were deemed not related to tonsillectomy (including two patients admitted to hospital with acute tonsillitis who had not undergone tonsillectomy).

### Discussion

Current guidance on adult tonsillectomy for recurrent acute tonsillitis is the same in the UK and the USA and is translated from the evidence for paediatric tonsillectomy rather than being based on adult-specific clinical trials.<sup>25</sup> NATTINA has shown that tonsillectomy is clinically effective for adults with recurrent tonsillitis meeting the UK guidelines.<sup>8</sup> Over the range of values that society is willing to pay for an additional QALY (assuming a

£10 000 threshold value), adult tonsillectomy has 100% probability of being considered cost-effective compared with conservative management.

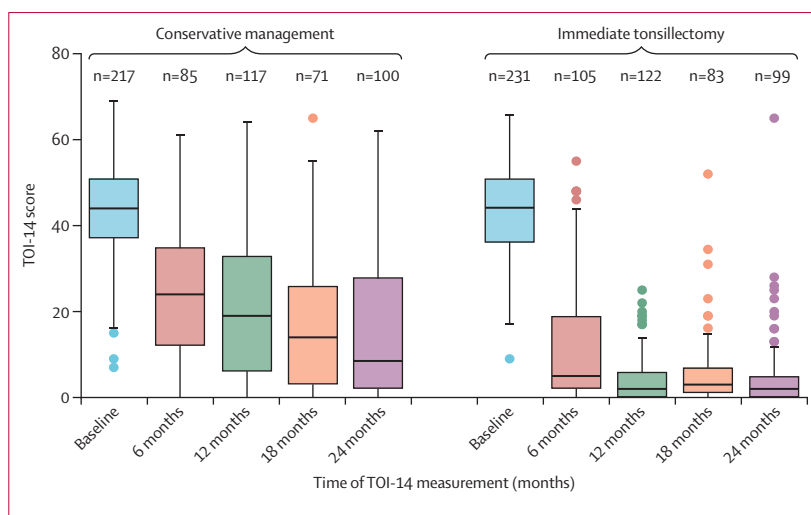
To our knowledge, NATTINA is the largest clinical trial to assess the effect of tonsillectomy in adults and offers clinically important long-term follow-up data. The benefit of immediate tonsillectomy, in terms of the 0.53 incidence rate ratio for sore throat days during 24 months versus standard conservative management, is clinically important. Patients considering tonsillectomy for recurrent acute tonsillitis should weigh the benefits of fewer sore throat days against the risks of surgery. NATTINA adds to the evidence to aid patients in shared decision making.

The postoperative bleeding rate of 19% is higher than reported previously,<sup>26</sup> but reflects the two proactive postoperative contact options actively made to elicit this information. The greater number of female participants recruited to NATTINA is consistent with the 70% adult female distribution undergoing tonsillectomy, which was shown in a large audit done in the UK.<sup>27</sup>

NATTINA materially augments the level 1 adult tonsillectomy evidence base. The 2014 Cochrane Review identified two studies showing that the number of sore throat days was 10.6 days fewer in those receiving tonsillectomy than those treated conservatively during approximately 6 months of follow-up. However, unlike our trial, neither study accounted for postoperative sore throat days. The anticipated 14.0 days of pain after a tonsillectomy needs to be considered by patients and health-care providers during shared decision making.

Despite existing adult tonsillectomy guidance, variation in practice regarding referrals remains, with many patients in the UK having difficulties accessing treatment. Douglas and colleagues<sup>28</sup> observed that the mean number of tonsillitis episodes before tonsillectomy was 27 during 7 years in 123 patients, which is three times higher than the UK guidelines of three episodes annually for 3 years. Findings from our study support adult tonsillectomy guidelines and should result in a practice change towards timely referral of patients.

Limitations of the study include the high dropout rates and missing primary outcome data (almost 10% higher in the conservative management group than the immediate tonsillectomy group when measured as those returning more than 80% once per week responses). The primary outcome of sore throat days was stipulated by the funder through a commissioned research call and recommended by the Cochrane Review as an appropriate measure. In a previous paediatric tonsillectomy trial<sup>14</sup> we showed that parents and patients found it challenging to recognise distinct episodes of tonsillitis as one bout merged with another, hence the decision to specify sore throat days rather than episodes of tonsillitis. Collection of data on a once per week basis during 104 weeks was challenging, but the continuous data collection approach minimised recall bias. Despite this challenge, 25% of participants completed every week and



**Figure 3: TOI-14 at baseline, and 6, 12, 18, and 24 months**

Box plot lines show the median scores and boxes show the lower and upper quartiles. The whiskers are 1.5 times the IQR. The points are outliers outside of that range. TOI-14=Tonsil Outcome Inventory-14.

	Number of episodes	Participants undergoing tonsillectomy (n=231)
Serious adverse events	51	45 (19%)
Bleeding	37	34 (15%)
Infection	6	6 (3%)
Pain	5	5 (2%)
Extended in-hospital stay	3	2 (1%)
Adverse events	191	90 (39%)
Bleeding*	54	44 (19%)
Returned to the operating theatre*	0	0
Readmission to hospital (all causes)*	47	47 (20%)
Readmission to hospital (bleeding)*	37	37 (16%)

Showing 172 patients assigned to immediate tonsillectomy plus 59 who requested switch from conservative management. \*Indicates postoperative complications.

**Table 3: Intervention-related safety outcomes**

57% responded in more than 80% of the 104 weeks. Further limitations include an absence of pain comparison between sore throat days and the postoperative period, and, although cited in previous studies, the TOI-14 questionnaire has not been formally validated in English from the original version in German.<sup>10</sup>

The NATTINA statistical analysis accounted for missing primary outcome data. The planned and unplanned sensitivity analyses support the main trial findings and conclusion. The unplanned sensitivity analyses show that the primary ITT results were maintained when restricted to participants returning more than 80% once per week sore throat responses. Because we assumed that any missing returns equated to no sore throat days, this sensitivity analysis offers the most conservative estimate of the effect of tonsillectomy.

The ITT analysis is probably an underestimate of the true impact of tonsillectomy in reducing sore throat days due to participant crossover to receive tonsillectomy. The ITT is also a reflection of real-world medicine with patients changing management options, particularly those waiting longer for a tonsillectomy. As expected, given the effect of tonsillectomy on sore throat frequency, the per protocol analysis (excluding crossovers) showed results more in favour of tonsillectomy than shown in the ITT analysis (incident rate ratio 0.42).

Although the economic evaluation was conducted in the UK, given that the management of recurrent acute tonsillitis is likely to be similar internationally and the rigorous methods adopted, these results still provide value for other countries. To facilitate decision making, data on surgical treatment received and health-care resource use are provided in the appendix (pp 15–20) to allow judgement of the generalisability of results and assign country-specific costs, if needed. Additionally, total costs were converted to US\$ and € and cost-effectiveness acceptability curves were derived to show the probability of tonsillectomy being considered as cost-effective with different threshold values for each currency.

Randomisation and continued follow-up is difficult in patients with recurrent tonsillitis. There are well documented challenges with recruiting participants to pragmatic trials, which aim to assess real-world effectiveness, especially when a surgical intervention is compared against a non-surgical intervention.<sup>29</sup> Patients eligible to enter NATTINA were referred to secondary care for consideration of tonsillectomy, having met the guidance level of tonsillitis episodes, and many had preconceived treatment wishes. Patients who declined entry had substantially worse recurrent tonsillitis symptoms, as measured by the TOI-14, as expected (appendix p 2). The patient population in this study might not represent those with worst outcomes, which could imply that the true benefit is greater than shown in this trial. Overall, we showed that compared with conservative management, immediate tonsillectomy is clinically effective and cost-effective in adults with recurrent acute tonsillitis.

#### Contributors

All authors wrote, reviewed, and revised the original manuscript. JAW, JO'H, DDS, LV, CHa, NR, SW, SC, CHo, FS, and AvW-M conceptualised the study. TF, DDS, MDT, LV, TH, and CHa did the formal analysis. JAW, JO'H, DDS, LV, SW, and CHa acquired funding. JAW, JO'H, TF, DDS, MDT, LV, FS, CHa, and MH contributed to the methods. JAW, CHa, JM, MDT, and DDS supervised the study. RW, AvW-M, and JAW were responsible for project administration. JAW, JO'H, CHa, KA-S, MH, SC, CHo, HM, and CR coordinated the trial and recruited participants. NH curated the data. All authors had full access to all the data in the study and accept responsibility to submit for publication. TF and MDT have accessed and verified all the data in the study.

#### Declaration of interests

DDS reports grants or contracts from National Institute for Health Research (NIHR), Yorkshire Cancer Research, Innovate UK, Heart Research UK, and Cancer Research UK; payment or honoraria from University of Liverpool and University of Leicester; support for attending meetings or travel (or both) from AntiCancer Fund and Royal College of

Surgeons of England; and participation on a data safety monitoring board or advisory board as a statistician and chair for NIHR. JAW reports support for attending meetings or travel (or both) from ENT Scotland. SC reports a leadership or fiduciary role for ENT UK. All other authors declare no competing interests.

#### Data sharing

Requests for de-identified data should be directed to the corresponding author. Data requests from researchers who provide a methodologically sound proposal will be considered up to 36 months from the date of publication.

#### Acknowledgments

This study was funded by the NIHR Health Technology Assessment programme (12/146/06). We thank the recruiting centres (Aneurin Bevan University Health Board; Bradford Teaching Hospitals NHS Foundation Trust; City Hospital Sunderland NHS Foundation Trust [South Tyneside and Sunderland NHS Foundation Trust]; Dorset County Hospital NHS Foundation Trust; East and North Hertfordshire NHS Trust; East Lancashire Hospitals NHS Trust [Royal Blackburn Hospital]; East Lancashire Hospitals NHS Trust [Royal Preston Hospital]; Frimley Health NHS Foundation Trust; Guy's and St Thomas's NHS Foundation Trust; James Paget University Hospital NHS Foundation Trust; NHS Ayrshire and Arran; NHS Grampian; NHS Greater Glasgow and Clyde; NHS Tayside; North Cumbria University Hospitals NHS Trust Cumberland Infirmary [North Cumbria Integrated Care NHS Foundation Trust]; Nottingham University Hospitals NHS Trust; Oxford University Hospitals NHS Foundation Trust; Plymouth Hospitals NHS Trust; Salisbury NHS Foundation Trust; Sheffield Teaching Hospitals NHS Foundation Trust; South Tees Hospitals NHS Foundation Trust; The Newcastle Upon Tyne Hospitals NHS Foundation Trust; United Lincolnshire Hospitals NHS Trust; University Hospitals Birmingham NHS Foundation Trust; West Suffolk NHS Foundation Trust; Worcestershire Acute Hospitals NHS Trust; and Wrightington, Wigan, and Leigh NHS Foundation Trust); Newcastle Clinical Trials Unit staff (including trial managers Isabel Rubie, Nicola Goudie, Faye Wolstenhulme, Rebecca Forbes, and Georgiana Browne; the clinical trial administrator Julia Phillipson; secretary Janet Jobling; the database officer Andy Cutts; and senior trial managers Lesley Hall, Chris Speed, and Jared Thornton); the health economist Peter McMeekin (study co-applicant); and statistician Shaun Hiu; and members of the Patient and Public Involvement group for advice and guidance when reviewing the trial and patient-facing documents. We also thank every participant involved in this trial; members of the trial steering committee (including John Birchall [chair]; Victoria Ward [independent clinician]; Catherine Hewitt [independent statistician]); independent lay members; and the data monitoring committee (Andrew Swift [chair] and Tim Woolford [independent clinician]).

#### References

- 1 National Health Service. Hospital episode statistics, admitted patient care England, 2009–10. 2010. <https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/hospital-episode-statistics-admitted-patient-care-england-2009-10> (accessed Dec 1, 2020).
- 2 Hall MJ, Schwartzman A, Zhang J, Liu X. Ambulatory surgery data from hospitals and ambulatory surgery centers: United States, 2010. *Natl Health Stat Rep* 2017; **102**: 1–15.
- 3 Burns RM, Wolstenholme J, Jawad S, et al. Economic analysis of oral dexamethasone for symptom relief of sore throat: the UK TOAST study. *BMJ Open* 2018; **8**: e019184.
- 4 Douglas CM, Altmyer U, Cottom L, Young D, Redding P, Clark LJ. A 20-year observational cohort of a 5 million patient population-tonsillectomy rates in the context of two national policy changes. *Clin Otolaryngol* 2019; **44**: 7–13.
- 5 Burton MJ, Glasziou PP, Chong LY, Venekamp RP. Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis. *Cochrane Database Syst Rev* 2014; **2014**: CD001802.
- 6 Chidambaram A, Nigam A, Cardozo AA. Anticipated absence from work ('sick leave') following routine ENT surgery: are we giving the correct advice? A postal questionnaire survey. *Clin Otolaryngol Allied Sci* 2001; **26**: 104–08.



- 7 Crowson MG, Ryan MA, Rocke DJ, Raynor EM, Puscas L. Variation in tonsillectomy rates by health care system type. *Int J Pediatr Otorhinolaryngol* 2017; **94**: 40–44.
- 8 ENTUK. Commissioning guide: tonsillectomy. 2021. [https://www.entuk.org/\\_userfiles/pages/files/guidelines/Revised%20ENT%20UK%20Tonsillectomy%20commissioning%20guide%20edit%20to%20final%20\(002\).pdf](https://www.entuk.org/_userfiles/pages/files/guidelines/Revised%20ENT%20UK%20Tonsillectomy%20commissioning%20guide%20edit%20to%20final%20(002).pdf) (accessed April 19, 2023).
- 9 Rubie I, Haighton C, O'Hara J, et al. The national randomised controlled trial of tonsillectomy in adults (NATTINA): a clinical and cost-effectiveness study: study protocol for a randomised control trial. *Trials* 2015; **16**: 263.
- 10 Skevas T, Klingmann C, Plinkert PK, Baumann I. Development and validation of the tonsillectomy outcome inventory 14. *HNO* 2012; **60**: 801–06 (in German).
- 11 Ware J Jr, Kosinski M, Keller SDA. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; **34**: 220–33.
- 12 McSweeney LA, Rousseau NS, Wilson JA, Wilkes S, Haighton CA. Stakeholders' views of recurrent sore throat, tonsillitis and their management: a qualitative interview study for the national trial of tonsillectomy in adults (NATTINA Part 1). *Clin Otolaryngol* 2017; **42**: 301–06.
- 13 Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004; **42**: 851–59.
- 14 Lock C, Wilson J, Steen N, 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 Technol Assess* 2010; **14**: 1–164, iii–iv.
- 15 Personal Social Services Research Unit. Unit costs of health and social care 2019. 2019. <https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/> (accessed Dec 1, 2021).
- 16 Matthews JN, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990; **300**: 230–35.
- 17 National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. 2022. <https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation> (accessed Feb 1, 2022).
- 18 Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res* 2011; **20**: 40–49.
- 19 Baltagi BH. Seemingly unrelated regression. In: Atkinson P, Delamont S, Cernat A, Sakshaug JW, Williams RA, eds. A companion to theoretical econometrics. Blackwell Publishing, 2003: 101–21.
- 20 Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Stat Med* 2000; **19**: 3219–36.
- 21 Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990; **10**: 212–14.
- 22 Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001; **10**: 779–87.
- 23 The Campbell and Cochrane Economics Methods Group and the Evidence for Policy and Practice Information and Coordinating Centre. CCEMG-EPPI-centre cost converter version 1.6. 2019. <https://eppi.ioe.ac.uk/costconversion/> (accessed Nov 11, 2022).
- 24 Husereau D, Drummond M, Augustovski F, et al. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *BJOG* 2022; **129**: 336–44.
- 25 Randall DA. Current indications for tonsillectomy and adenoidectomy. *J Am Board Fam Med* 2020; **33**: 1025–30.
- 26 Lowe D, van der Meulen J. Tonsillectomy technique as a risk factor for postoperative haemorrhage. *Lancet* 2004; **364**: 697–702.
- 27 The Royal College of Surgeons of England. National prospective tonsillectomy audit. 2005. <https://www.rcseng.ac.uk/library-and-publications/rcs-publications/docs/tonsillectomy-audit/> (accessed Dec 11, 2021).
- 28 Douglas CM, Lang K, Whitmer WM, Wilson JA, Mackenzie K. The effect of tonsillectomy on the morbidity from recurrent tonsillitis. *Clin Otolaryngol* 2017; **42**: 1206–10.
- 29 Cook JA. The challenges faced in the design, conduct and analysis of surgical randomised controlled trials. *Trials* 2009; **10**: 9.